Received: 26 November 2020 | Revised: 28 January 2021 | Accepted: 22 February 2021

DOI: 10.1002/ett.4253

RESEARCH ARTICLE

WILEY

Minimum energy transmission forest-based geocast in software-defined wireless sensor networks

Anuradha Banerjee[1] | Abu Sufian[2] | Ali Safaa Sadiq[3] | Seyedali Mirjalili[4,5]

[1]Department of Computer Application, Kalyani Government Engineering College, Kalyani, India

[2]Department of Computer Science, University of Gour Banga, Malda, India

[3]Wolverhampton Cyber Research Institute, School of Mathematics and Computer Science, University of Wolverhampton, Wolverhampton, UK

[4]Centre for Artificial Intelligence Research and Optimisation, Torrens University Australia, Brisbane, Queensland, Australia

[5]Yonsei Frontier Lab, Yonsei University, Seoul, Republic of Korea

Correspondence
Ali Safaa Sadiq, Wolverhampton Cyber Research Institute, School of Mathematics and Computer Science, University of Wolverhampton, Wulfruna Street, Wolverhampton WV1 1LY, UK.
Email: ali.sadiq@wlv.ac.uk

Abstract

Wireless Sensor Networks (WSNs)-based geographic addressing and routing have many potential applications. Geocast protocols should be made energy efficient to increase the lifetime of nodes and packet delivery ratio. This technique will increase the number of live nodes, reduce message costs, and enhance network throughput. All geocast protocols in the literature of WSN apply mostly restricted flooding and perimeter flooding, which is why still the redundancy they produce significantly high message transmission costs and unnecessarily eats up immense energy of the nodes. Moreover, perimeter flooding cannot succeed in the presence of holes. The present article models WSN with software-defined (SD) constructs where the network area is divided into some zones. Energy-efficient transmission tree(s) are constructed in the geocast area to organize the flow of data packets and their links. Therefore, redundancy in the transmission is eliminated while maintaining network throughput as good as regular flooding. Besides, this study introduces a fuzzy controllers named SELECT-STARTER and GEOCAST-PROPAGATOR for intellgent energy tansmission control-making. The proposed technique significantly reduces energy cost and improves nodes' lifetime to function for higher time duration and produce a higher data packet delivery ratio. To the best of the author's knowledge, this is the first work on geocast in SD-WSNs.

1 | INTRODUCTION

These days' researchers have developed a great interest in Internet-of-Things. Wireless sensor networks or WSN form a big part of it.[1-10] These networks consist of specific nodes that communicate with each other in the single or multi-hop. In multiple hops, several nodes have to play the role of a router, and nodes deplete quickly after they participate in a large number of communication sessions, especially if routes of those sessions break frequently and require more than one route rediscovery.[10-22]

Geocast[1,7,23,24] is a special kind of broadcast operation where one particular region in the network is predefined, and all data packets are targeted to this location. All sensor nodes residing within this region can receive this message, while nodes outside this region should not receive it. Existing literature on geocasting in WSN applies two different flooding directions—directional flooding and local flooding. They are briefly described below in Figures 1 and 2, while Figure 3 introduces software-defined networks or SDN to illustrate the utility of SDN in the context of geocast.

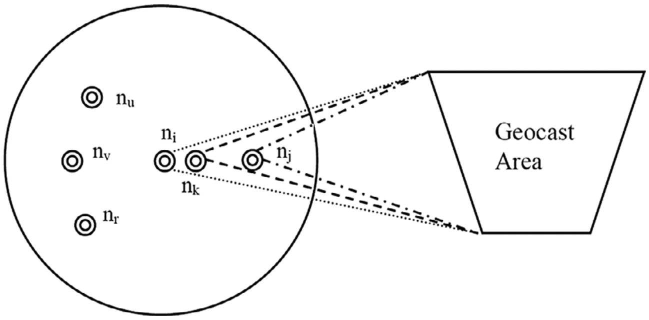

FIGURE 1 Directional flooding

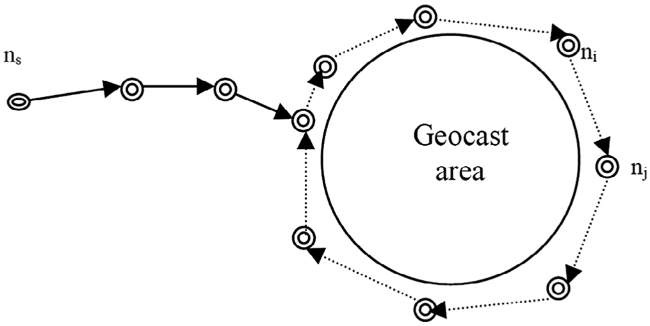

FIGURE 2 Perimeter flooding

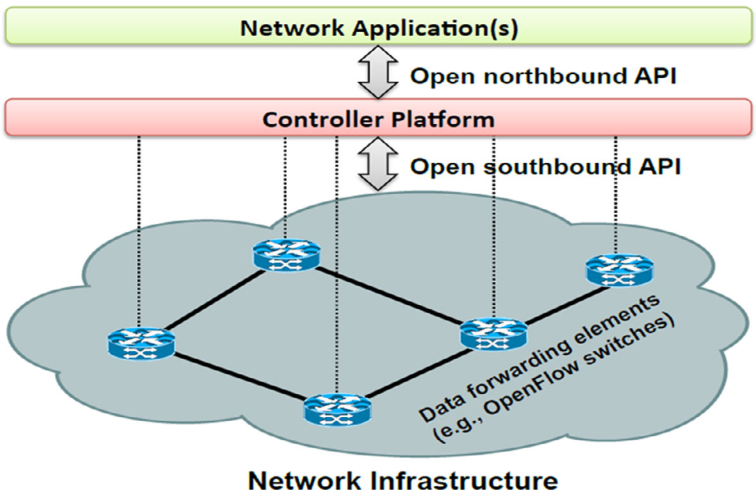

FIGURE 3 Structure of a software-defined networking

Studies[25,26] discuss several efficient and robust geocast routing protocol. They are, Geographic-Forwarding-Geocast (GFG) Geographic-Forwarding-Perimeter-Geocast (GFPG). GFG reduces cost in dense networks while GFPG guarantees geocast packet delivery but at some extra messages. These are particularly applicable for WSN but lose significance in the context of SD-WSN because, in SD-WSN, nodes are static, and delivery can be guaranteed only if nodes remain alive till the geocast message reaches it. There is no question of getting out of geocast region with time. Therefore, velocity consideration is insignificant. The zonal structure in our network has been motivated by ISFC-BLS (Intelligent and Secured Fuzzy Clustering Algorithm Using Balanced Load Sub-Cluster Formation) in WSN Environment[27] where the network is divided into some load-balanced clusters and ant colony optimization is applied to find optimal path to destination. It provides good energy efficiency compared to the ring and other clustering methods applicable for WSN. Specific protocols also bring energy efficiency into the system. For example, ISFC-BLS[27] deploys rechargeable sensors and dynamically balances load through the maximum capacity path or MCP technique. A fuzzy genetic-based dynamic spectrum allocation system for deciding spectrum allocation is proposed in References 28-30 where parameters like bit error rate, signal interference noise ratio, available channel bandwidth, and sender unit (SU) transmission power are monitored continuously so that performance of the system can be optimized by reducing cost. Hybrid Energy-Efficient Distributed clustering[31] periodically selects cluster heads depending on two parameters—residual energy and node degree. Therefore this clustering strategy is energy efficient as well. Despite numerous such protocols, the network's smooth functioning is often hampered by the attackers due to the open nature of WSN as described in References 32,33. Reference 31 discusses various potential threats at different protocol layers and various routing and middleware challenges in detail. Approaches like trust management, intrusion detection, firewalls, and key management are discussed in Reference 33. However, in the present context, this is not the targeted problem area. A survey is proposed in Reference 34 where the basics of SDN, WSN, and SD-WSN are explained. It describes the benefits that SDN architecture can incorporate in WSN. Mainly this article has motivated us to build our geocast protocol in SDWSN construct. An energy-efficient multicast routing protocol based on SDN and fog computing for vehicular networks is proposed in Reference 34 where multiple quality of service (QoS) constraints are satisfied. Also, a scheduling algorithm is proposed that prioritizes multicast packets based on application type and deadline. Another energy-efficient algorithm is proposed in Reference 35 for selecting the optimum cluster heads among all cluster members. Both delay and energy are considered as performance metrics in Reference 36.

A Congestion Aware Routing using Fuzzy Rule sets has been proposed in Reference 37. It applies fuzzy rule bases for identifying the nonlocalized node paths to add to the existing localized node paths and selecting the least congestion affected destination path. The proposed technique increases the lifetime of nodes and significantly decreases the percentage of packet drops due to energy scarcity or nodes' exhaustion. Multicast protocols can always be utilized in geocast. However, in the latter case, all receivers' identifiers and locations within the geocast region need to be supplied to the sender, which is difficult and assumed as unknown by conventional geocast protocols. The current work has been framed abiding by those constraints.

As far as geocasting in WSN is concerned, two popular protocols are directional and perimeter flooding. Directional flooding[5,8,18] instructs a node to forward data packets to only those downlink neighbors which are in the direction of geocast area. For example in Figure 3, it can be seen that node n_i has six neighbors n_j, n_k, n_l, n_u, n_v,, and n_r. Among them, only n_j and n_k are in the direction of geocast area while the others are in the opposite direction. Therefore, n_i sends geocast data packet to n_j and n_k only, for forwarding. After receiving the packet, n_j and n_k follow the same procedure. This saves some message cost and also reduces redundancy. For example, in Figure 1, assume that n_k and n_j are in the neighborhood of each other. After receiving geocast packet from n_i, n_j will not send the same to n_k because n_k is not in the direction of geocast area from the point of view of n_j. This is correct from the redundancy perspective too. n_k has already received the message from n_i and there is no point in receiving the same from n_j. But, as far as n_j is concerned, this will receive geocast packet from both n_i and n_k because n_j lies in the forwarding region of both, generating redundancy.

This is demonstrated in Figure 2. Source and each router forward the data packet to the node that is closest to an approximate center of the geocast region.[11,38,39] As soon as the data packet reaches a router that has at least one neighbor in the geocast region, the router starts forwarding the packet along the boundary of the geocast area, so that a perimeter of geocast area is constructed. After that all nodes residing in the constructed periphery, starts flooding the data packet to all neighbors inside the geocast region. Redundancy within the geocast area is unavoidable here.

Perimeter flooding works well in a densely populated environment otherwise the problem of the hole can wreck havoc with its functionality. A hole in the nodal periphery evolves when the radio-circle of a node in the periphery is completely empty. For example, in Figure 2, suppose node n_i has no neighbor to forward packet and n_i is also not within the neighborhood of n_j. Hence there will be a hole between n_i and n_j and construction of nodal periphery would not be incomplete. Without this periphery, flooding inside geocast area would not be possible.

Here the authors have proposed geocasting in SDN-based WSN. The main idea behind SDN[40-42] is to separate the physical forwarding of data and its control through software. Such networks are divided into three components—a centralized controller, some hosts, and some switches. Hosts can only issue flow requests, that is, they can inject flows into the network and receive packets from it. They do not have forwarding capacities. On the other hand, the sole purpose of a switch is to forward packets either from a host to another host or from a switch to a host, or from a host to a switch or in between two switches. All these forwarding activities are performed under the direction of one centralized controller. Communication between different entities is performed through APIs or application programming interfaces. Instructions of centralized controller come in the form of flow tables that are used for configuring forwarding planes. The basic structure of SDNs appears in Figure 3.

All three local, metropolitan and wide area networks apply SDN concepts. It is used in data centers as well, for flexible network control without compromising with forwarding references. This kind of network is easy to implement and maintain. Also lots of scopes are there to improve the energy efficiency of these protocols.[20,43] Some common methods are occasional deactivation of switches, reducing the size of content addressable memory, etc. SDN has successfully proved its worth and vast applicability in highly scalable data centers in the private and public sectors alike. SDNs have set a new trend with their performance effectiveness.[44] Nowadays hybrid SDN are gaining popularity for several benefits,[45] SD-WSN is one such network. In this proposed work, SDN divides the completely distributed structure of WSN into some nonoverlapping clusters or zones where each controller is in charge of one zone. Peripheral nodes in a zone are nodes that have some portion of the radio-circle in other zones. Those zones are termed as neighbors of the current zone. Our work proposes an energy-efficient green geocast protocol having certain novelties. To the best of authors' knowledge, the present work is the first geocast protocol in SD WSN. This is a combined model of SDN and WSN whereas its competitors are only geocast protocols in WSN. Normally in WSN, whenever a geocast operation is initiated by a source, it performs flooding with a few restrictions or tricks but still suffers from significant redundancies, especially when the network is densely populated. This unnecessarily eats up energy in nodes reducing their lifetime. This green geocast is capable of completely eliminating flooding and still achieving high network throughput, often as good as flooding. The proposed protocol constructs energy efficient transmission tree(s) within the portion of geocast region inside each zone. Here, the cost is a function that depends on energy, remaining lifetime, etc. Data packets are delivered along energy-efficient links of these trees to save energy in nodes. Still some nodes in geocast region of a zone may remain isolated. For them, energy-efficient unicast routes are formed from source (if the source is in the current zone) or the peripheral node through which geocast packet entered into the current zone (if the source is not in the current zone) to the isolated nodes. Forming energy efficient transmission tree as well as unicast routes to isolated fellows, are governed by fuzzy controllers embedded in the SDN controller of a zone.

The rest of the article is:related works appear in Section 2. Network model details are presented in Section 3, and Section 5 contains a simulation experiment and the analysis of results; the conclusion is in Section 6.

2 | RELATED WORK

Although this routing model is novel but some related state-of-the-art geocast protocols are mentioned the following.

2.1 | Minimum energy communication networks

In minimum energy communication networks (MECN),[46] the network is modeled as a graph consisting of certain vertices and edges. Each edge has got some cost associated with it. MECN applies Bellman-Ford shortest path algorithm to find out minimum cost path between each pair of vertices. In this process it eliminates inefficient edges from the original network graph to arrive at a smaller subnetwork graph where all vertices of the original network graph are present with only least cost edges connecting them. Definitely, network connectivity is preserved in the process, which is no two nodes lose connectivity during the elimination of high-cost links.

MECN works according to the inherent distributed architecture of WSN. Therefore, discovering an energy-efficient route requires the exchange of a huge number of route-request route-reply messages incurring significant cost. Moreover nodes regularly transmit hello messages within their respective radio-ranges to periodically detect its neighbors. This too eats up a lot of energy. MECN follows directional flooding strategy.

2.2 | Geographic adaptive fidelity

In geographic adaptive fidelity (GAF),[39] certain routers are allowed to take rest if an appropriate alternative is available. For example, if two nodes n_a and n_b have set of uplink downlink neighbors $\{n_c, n_e, n_f\}$ and same set of downlink neighbors $\{n_p, n_q, n_r\}$, then one can play others role when the other one is sleeping. For example, assume that a live communication session is going on in the path $\dots \rightarrow n_e \rightarrow n_a \rightarrow n_r \rightarrow \dots$ while n_a does not have much charge left and battery of n_b is running lively with huge remaining charge. Then the path $\dots \rightarrow n_e \rightarrow n_a \rightarrow n_r \rightarrow \dots$ will change to $\dots \rightarrow n_e \rightarrow n_b \rightarrow n_r \rightarrow \dots$ and n_a will be given the chance to sleep. To perform geocastin GAF, a virtual peripheral grid is constructed around the region of geocast. A node can either go to sleep (sleep state) or engage itself in route discovery task (discovery state) when nodes determine their numbers in the grid. Route discovery means exchanging a huge number of route-request route-reply messages. These messages, coupled with hello and ack or acknowledgment, consume huge energy in nodes. GAF maintains each active node as representative of each virtual grid, to keep the network connected. This active node bears the responsibility of flooding packets inside the grid.Flooding is directional in this protocol; directed toward a particular grid.

2.3 | Geographic energy aware routing

Geographic energy aware routing (GEAR)[47] is an energy efficient geocast protocol emphasizing on the fact that there are two tasks in an geocast—first one is to arrive at the geocast region and second one is to broadcast geocast message in that region. Whenever a node outside geocast region receives an geocast message for forwarding, it chooses a neighbor in its list of neighbors that is at minimum distance from centroid of geocast region. Packets are directed toward geocast region, hence method applied here is directional flooding. This protocol does not differentiate between two hops that consume different energy owing to minimum receive power of receivers in those hops.

2.4 | Energy aware geographic routing with anchor nodes

In energy aware geographic routing (EAGR),[48] energy cost corresponding to each forwarding decision, is computed in advance. Each sender adjusts its transmission power so that it is just enough to reach its next hop receiver. But if receiver is on radio-range of the sender then no reduction in transmission power will be possible. Geocasting is performed using directional flooding method.

2.5 | Geographic routing based on partitioned architecture

This is a work that considers multiple static sinks[8] and protection of their information from one another. Nodes involved with multiple sinks are divided into several levels. Sinks that logically appear in more than one levels, receive real packets while fake packets are forwarded to fake sinks through multiple unicasting. It does not use flooding. The idea is to improve robustness of the protocol and make it efficient enough to handle changes in topology with more than one sinks, without imposing much communication load and energy consumption.

2.6 | Greedy parameter stateless routing

Like MECN, greedy parameter stateless routing (GPSR)[16] also models the underlying network as a graph and produces a simpler planar graph out of it. Planar graph is the one used for propagation of data packets. But one thing that can harm construction of perimeter (perimeter flooding is applied in this protocol) around geocast zone, is presence of holes. Its number may increase in between a communication session because nodes in the perimeter tend to deplete quickly due to load of flooding in geocast region. This runs the risk of breaking wall of perimeter producing more holes.

2.7 | Energy efficient location-based intersection routing

Energy efficient location-based intersection routing (EELIR)[18] is based on the concept that each node knows its residual energy and current location. Priority of a node increases with increase in its residual energy and optimal route is the one that contains comparatively highest number of high priority nodes. Geocast source discovers routes to individual geocast members through directional flooding. Therefore flooding is required only once for route discovery. During delivery of data packets no flooding is required. This solves the problem of redundancy in flooding. Message cost in EELIR will be smaller than others if geocast source is close to the geocast region and geocast region is not very densely populated.

2.8 | Anonymous location-based energy-efficient routing

Anonymous location-based energy-efficient routing [22] is distinguished by its low cost and anonymity protection for sources, destinations, and routes. It assumes that entire network area is a rectangle and uses recursive hierarchical as well as vertical partition until desired zone is obtained. After that, in each step, a forwarding node is randomly chosen which is closest to the zone. In the last step, data packets are broadcast to destination zone. Broadcasting is performed using directional flooding technique.

2.9 | Fixed rectangular forwarding zone geocast

In fixed rectangular forwarding zone geocast (FRFZ),[26] the sender computes vertices of the smallest rectangle that circumscribes the geocast region, from knowledge of simple geometry. As soon as it reaches a node which is close to the vertex, that is, distance between the vertex and the node is less than a predefined threshold, then sender sends geocast message to that node. Similarly other three nodes are discovered and all of them receive geocast message from the sender. Then these nodes flood the same to the geocast region. If a node is inside the geocast region, it processes the message, otherwise the message is dropped.

3 | NETWORK MODEL DETAILS

In this section all details of the current proposed model are explained categorically. Each issue are described in following subsections. The SD-WSN framework consists of the following layers as shown in Figure 4.

(i) **Physical layer (PL)**: Multiple sensor nodes exist at this level. They communicate with the SDN controller using Openflow[6] protocol to receive instruction in the form of flow table information. Flows are directed according to it. Please note that nodes are divided into certain zones and each zone is under supervision of exactly one controller. Please note that, improvement in operation of SD-WSN depends on three aspects—suitable controller placement and re-election,[49] efficient management of Openflow switches,[50] and proper control of data plane.[51]

(ii) **Virtualization layer**: The key nodes of the sensor network in the PL map to virtual key nodes, in order to form the virtual node layer (virtual layer).

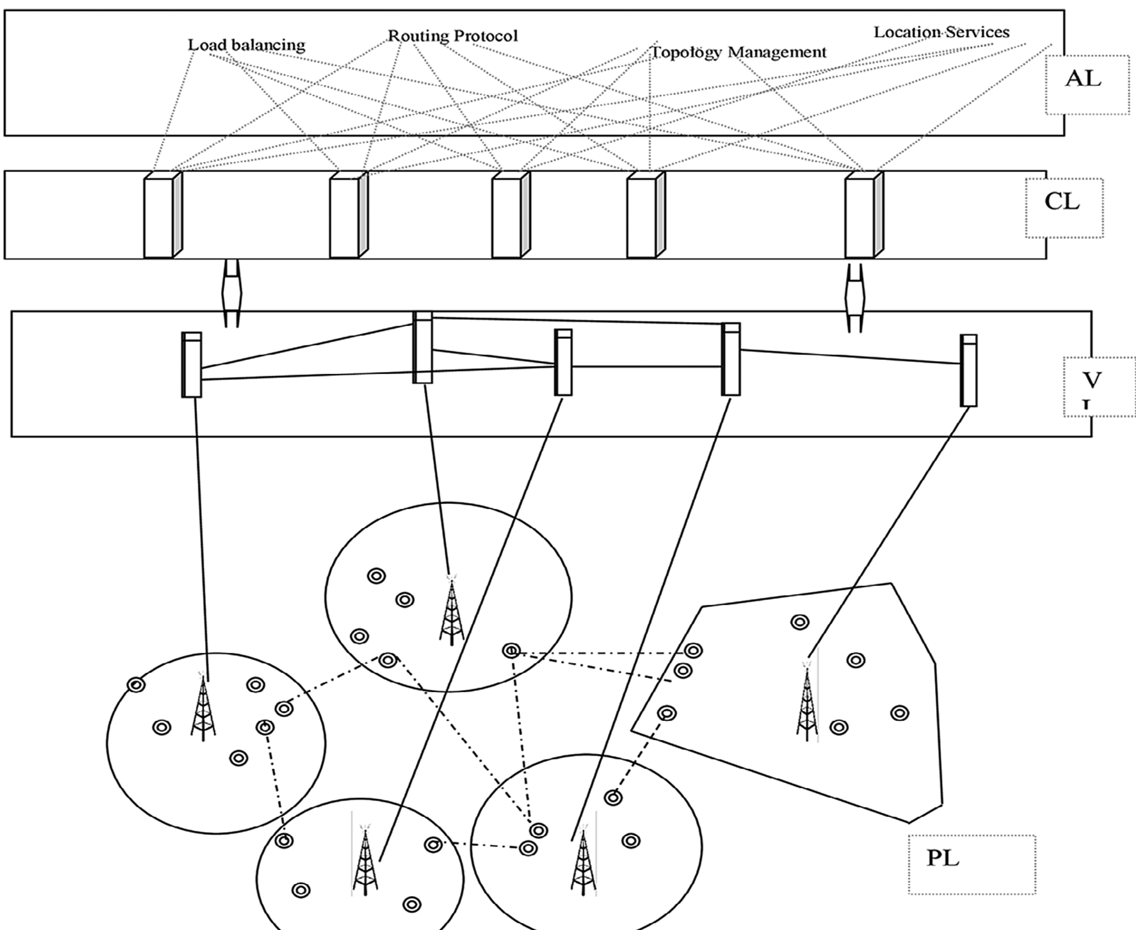

FIGURE 4 Network structure

(iii) **Control layer (CL)**: This layer consists of SDN controllers which control routing, sleep monitoring, and other services which are components of network management. It communicates its flow control instructions to individual nodes using Openflow protocol. Based on location, radio-ranges, battery power, etc. information, an optimum route is selected from one node in the current zone to some other node (maybe in a same or different zone) an optimal route is selected.
(iv) **Application layer (AL):** The AL defines a working policy for each application. Taking decisions about working strategy requires information about network topology and node status which are available from the CL. These strategies decide how services will be provided to sensor nodes.

3.1 | Information stored in SDN controller

SDN controller of a zone Z_p ($1 \leq p \leq P$; P is the number of all zones in the network) consists of the two following tables,

(i) TOPO-INFO$_p$
(ii) NODE-INFO$_p$

TOPO-INFO$_p$ specifies topology of zone Z_p with its attributes being

(i) node-id
(ii) adjacency-list

node-id specifies unique identification number of a node whereas adjacency-list is the set of all 1-hop downlink neighbors of the node.

NODE-INFO$_p$ is a table consisting of following information about nodes in Z_p.

(i) node-id
(ii) (latitude, longitude) ordered pair
(iii) radio-range
(iv) maximum energy
(v) most recent residual energy
(vi) most recent rate of energy depletion
(vii) minimum receive power
(viii) peripheral status
(ix) current time

The latitude-longitudinal pair of a node n_i isrepresented by $(x_i(t), y_i(t))$ while its radio-range is rad$_i$. The highest battery capacity of n_i is m-en$_i$ while its residual energy at time t is r-en$_i(t)$ and current rate of energy depletion is given by dep$_i(t)$.The peripheral status of n_i is pphr$_i$. It is set to 1 if it has some neighbors belonging to any neighboring zone of Z_p. Otherwise it is set to 0.

As soon as a node n_i starts operating in a location under zone Z_p, it sends a register message to SDN controller. Attributes of a register message are,

(i) msg-id (set to 0)
(ii) node-id
(iii) (latitude, longitude) ordered pair
(iv) radio-range
(v) highest battery power

After a register message is received by the SDN controller of zone Z_p, it searches NODE-INFO$_p$ table to check if it is a duplicate register message or not. If an entry with the same identification number already exists in the table, then the SDN controller replies with a duplicate-entry message. Attributes of a duplicate-entry message are-.

(i) msg-id (set to 1)
(ii) node-id

But if no such duplicate entry is found in NODE-INFO$_p$, then a success message is sent back to the node that generated register message. Its components are the same as a duplicate entry; only msg-id is assigned the value 2. The next task is to determine the peripheral status of n_i. It broadcasts a hello message within its own radio-range. Identification numbers and locations of all nodes that acknowledge this message are sent to the SDN controller. If at least one of those nodes is outside the current zone, then SDN will set pphr$_i$ to 1; otherwise it will be set to 0. Attributes of hello and ack messages are as follows:

(i) msg-id
(ii) node-id
(iii) (latitude, longitude) ordered pair
(iv) radio-range

The msg-id is 3 for hello and 4 for ack. Three types of zone structures are considered—polygonal, circular, and elliptic. Similar are the structures of geocast regions. Polygonal structure is closest to the real-life scenario while circular is the ideal one where neighborhood of a node spreads around the node equally in all directions. Please note that circle is a special kind of ellipse where both foci are at the same point (the center). Therefore both circular and elliptic structure has been considered in our simulation. Below the authors derive the conditions in section V under which a polygonal, circular or elliptic geocast region embeds a zone member n_i. Before beginning the geocast operation, an geocast-initiation message is sent. Its attributes are:

(i) geocast source-id
(ii) zone-id of geocast source-id
(iii) geocast area specification

After a zone identifies that it has at least one geocast member in it (the process is illustrated in Section 4.1), it sends a reply named geocast-reply to the geocast source. Components of geocast reply are.

(i) zone-id
(ii) List of geocast member id and their locations

After sending geocast-initiation message, geocast source waits for time τ-reloc and then starts sending data packets.

4 | OVERVIEW OF THE PROPOSED WORK

In the beginning, immediately after receiving a geocast request, the SDN controller checks to see if the geographical region circumscribes the entire zone. If the whole zone falls within the geographic region, then all nodes in the zone are intended geocast receivers. Otherwise only a subset of nodes will be geocast receivers. Among these nodes, certain nodes have to be elected as starter and role of the starter is to propagate geocast message to all nodes reachable from it. There may be multiple starters in the geocast region if all nodes are not reachable from the current starter. After that energy efficient transmission tree corresponding to each starter is constructed so that the starters can parallely proceed with the geocast transmission. Lonely leaves can be elevated as much as possible so that delay in transmission can be reduced. Figure 5 shows these steps in a schematic diagram.

4.1 | Geocast members

All nodes in a zone may not be geocast members. SDN controller determines whether a node is geocast member or not depending upon specification of the area of geocast region. If geocast region is a circle, then geocast source/peripheral node informs SDN controller about its center and radius. Similarly if it is circular then center and foci lengths are specified and in case of polygonal geocast area, sequence of consecutive edges starting from one particular vertex, are specified.

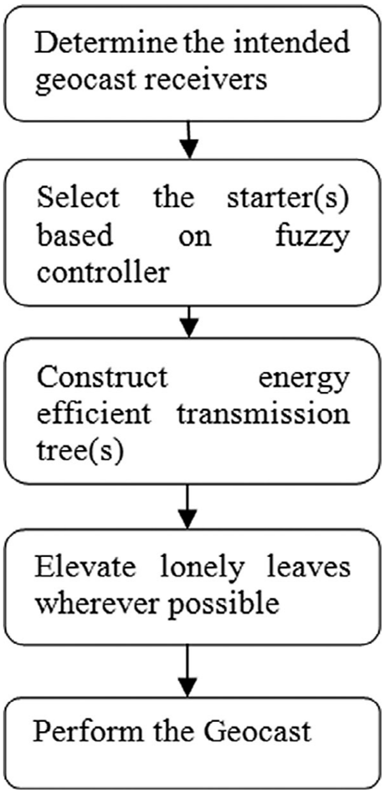

FIGURE 5 Diagram of the proposed work

Position of n_i at current time t, is denoted as the ordered pair $(x_i(t), y_i(t))$. With respect to a circular zone defined by center (h, k) and radius r, n_i will be inside provided $f_1(n_i,t) < 0$ where $f_1(n_i,t)$ is expressed in (1).

$$f_1(n_i, t) = (x_i(t) - h)^2 + (y_i(t) - k)^2 - r^2. \tag{1}$$

With respect to an elliptic zone defined by center (h, k) and major and minor axis length a and b, n_i will be inside at current time t provided $f_2(n_i,t) < 0$ where $f_2(n_i,t)$ is mathematically expressed in (2).

$$f_2(n_i, t) = \{(x_i(t) - h)/a\}^2 + \{(y_i(t) - k)/b\}^2 - 1. \tag{2}$$

Situation becomes just a bit more complex when geocast region is polygonal defined by the vertices $(x_v, y_v), (x_{v+1}, y_{v+1}), (x_{v+2}, y_{v+2}), \ldots, (x_{v+j}, y_{v+j})$.. Let,

$$\left.\begin{aligned} x_{\min} &= \forall \min\,(x_{v+m}), \text{s.t.} \quad m = 0 \text{ to } j \\ y_{\min} &= \forall \min\,(y_{v+m}), \text{s.t.} \quad m = 0 \text{ to } j \end{aligned}\right\} \tag{3}$$

n_i will be completely outside the current zone at present timestamp t, if $x_i(t) < x_{\min}$ or $x_i(t) > x_{\max}$ or $y_i(t) < y_{\min}$ or $y_i(t) > y_{\max}$. Otherwise, n_i is treated as a member candidate. In order to determine whether n_i is a node inside the zone, a straight line parallel to X-axis is drawn that passes through $(x_i(t), y_i(t))$ of that line is $y = y_i(t)$. This line will intersect a polygonal edge from $(x_{v+k'}, y_{v+k'})$ to $(x_{v+k'+1}, y_{v+k'+1})$ where $0 \le k' \le (j-1)$ provided $y_{v+k'} \le y_i(t) \le y_{v+k'+1}$. Coordinate of the intersection point is $(f_i(v, k', t), y_i(t))$ where $f_i(v, k', t)$ is mathematically expressed in (4).

$$f_i(v, k', t) = x_{v+k'} + (y_i(t) - y_{v+k'})(x_{v+k'+1} - x_{v+k'})/f_n(v, k') \tag{4}$$

where $f_n(v, k') = (y_{v+k'+1} - y_{v+k'})$.

The intersection points are arranged from left to right in increasing order of x-coordinates and then duplicates are eliminated. Please note that, duplicates, in this case, appear consecutively, if they appear at all. Let intersection points are given by $(a_q, y_i(t))$ such that $a_q \le a_{q+1}$, where q is an integer between 0 and j. Among them if there are any two intersection points $(a_q, y_i(t))$ and $(a_{q+1}, y_i(t))$ where $a_q = a_{q+1,}$ then $(a_{q+1}, y_i(t))$ is removed from the set of intersection points. Assume that after eliminating duplicates, remaining number of intersection points is *c*.

In Figure 6A it can be seen that ABCDE is a polygon where coordinates of A, B, C, D and E are (x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4), and (x_5, y_5). For the position of n_i in Figure 6A, the straight line $y = y_i(t)$ intersects polygonal edges AB, AF, EF and DE at points P, Q, R and S with their coordinates being $\{x_1 + (y_i(t) - y_1)(x_1\text{-}x_2)/(y_1 - y_2), y_i(t)\}$, $\{x_1 + (y_i(t) - y_1)(x_1 - x_6)/(y_1 - y_6), y_i(t)$, $\{x_5 + (y_i(t) - y_5)(x_5 - x_6)/(y_5 - y_6), y_i(t)\}$ and $\{x_5 + (y_i(t) - y_5)(x_5 - x_4)/(y_5 - y_4), y_i(t)\}$, respectively. Therefore, if n_i has to exist within the polygon then it will have to exist on the straight line $y = y_i(t)$, in between *z*th odd intersection point and $z+1$-th intersection point where $1 \leq z \leq |\text{PT-ODD}|$. Please note that PT-ODD is the set of all odd intersection points, after arranging them from left to right, that is, in increasing order of x coordinates. With every odd intersection point, direction of movement is towards interior of the polygon whereas with every even intersection point, direction of movement is toward exterior of the polygon. But what will happen if number of intersection points is odd? This is illustrated in Figure 5B.

It may be noticed in Figure 6B that, three intersection points are there. A is obtained twice—once for intersection of the straight line $y = y_i(t)$with AB and then with AF. Another intersection point is *Q* which is obtained when the same line $y = y_i(t)$ cuts polygon edge BC. After arranging these intersection points from left to right, the obtained sequence is (A, A, Q). After eliminating consecutive duplicates, the new sequence is (A, Q). If n_i lies between A and Q, that is, if $x_i(t)$ is in between (x_1, y_1) and $\{x_2 + (y_i(t) - y_2)(x_2 - x_3)/(y_2 - y_3)\}$, then it belongs to the polygon ABCDEF; otherwise it is outside.

4.2 | Message transmission in green geocast

For transmitting geocast messages to all nodes belonging to a portion of geocast region inside the current zone, a minimum energy transmission forest is constructed. An energy-efficient transmission forest FR is a collection of minimum energy transmission trees denoted by TR(*y*) as in (5) such that $1 \leq y \leq |\phi|$ where ϕ is construction of minimum energy transmission forest for a set of nodes V has to begin with one particular starter node. Starter node is one that receives geocast message directly from the geocast source (if geocast source resides in the current zone), or a peripheral node

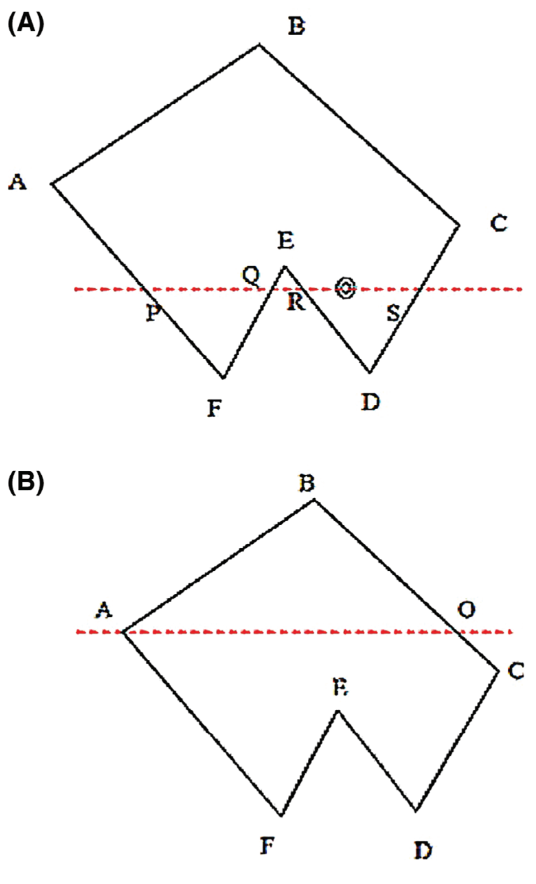

FIGURE 6 (A) Number of intersection points is even (P, Q, R, S). (B) Number of intersection points is odd (A, A, Q)

(in the situation when geocast source belongs to any of the remaining zones). Role of the starter is to propagate geocast message to all nodes reachable from it. There may be multiple starters in the geocast region if all nodes are not reachable from the current starter. If all nodes are reachable from the selected starter, the minimum energy transmission forest will consist of only one minimum energy transmission tree; otherwise multiple such trees will be there. Each tree will have exactly one starter node. It is an geocast member that cannot have any geocast other member as uplink neighbor in minimum energy transmission tree.

A minimum energy transmission tree $\text{TR}(y) = <\text{Vr}(y), \text{Ed}(y), \text{Wg}(y,t)>$ such that $1 \le y \le |\phi|$, $\text{Vr}(y)$ is the set of vertices, $\text{Ed}(y)$ is the set of edges, and $\text{Wg}(y,t)$ is the summation of weights of all edges at time t, as formulated in (6). Let each edge e_j: $(n_k \to n_p)$ s.t. $1 \le j \le |\text{Ed}(y)|$ and $n_p, n_v \in \text{Vr}(y)$ has an weight $w_j(t)$ at current timestamp t. Based on this, $\text{Wg}(y,t)$ is expressed below.

$$\text{Wg}(y, t) = \Sigma w_j(t), \text{s.t.}, \ j = 1 \text{ to } |\text{Ed}(y)| \quad . \tag{6}$$

In proposed geocast algorithm, weight $w_j(t)$ of an edge e_j at time t is directly proportional to expected energy consumption along the link from n_k to n_p. Therefore, $w_j(t)$ is mathematically formulated in (7) based on Frii's transmission for estimation of energy consumption in wireless links.

$$w_j(t) = \{\text{mrp}(p)\ \text{dist}_t^{\ 2}(k, p)\}/\text{C1}. \tag{7}$$

$\text{mrp}(p)$ is minimum receive power of n_p. C1 is a constant.

Delay $d_j(t)$ of the edge e_j is the time duration required for sending one packet from n_k to n_p at time t. It is modeled in (8).

$$d_j(t) = \text{C2}\ \text{dist}_t(k, p) \quad . \tag{8}$$

C2 is another constant.

Among two links $e_j : (n_k \to n_p)$ and $e_r : (n_u \to n_v)$, e_j will be considered more efficient if either condition CN1 or CN2 is true. CN1 specifies that if e_j is more energy efficient, then irrespective of delay, it is considered better than e_r. CN2 is applicable only when energy consumption from n_k to n_p is same as energy consumption from n_u to n_v, that is, $w_j = w_r$.

$$\left.\begin{array}{l} \text{CN1} : w_j < w_r \\ \text{CN2} : d_j < d_r \end{array}\right\} .$$

The present article aims at constructing energy efficient transmission forest for geocasting packets to all geocast members in the current zone and then choosing suitable peripheral nodes to propagate to some / all of the neighbor zones if required. Selection of starter and selecting neighbor nodes for propagation of geocast message to neighbor zones, are all governed by fuzzy controllers named SELECT-STARTER and GEOCAST-PROPAGATOR respectively. Pseudocode of determining whether a node is inside geocast region or not are explain in Figure 7.

4.3 | Construction of energy efficient transmission forest

The selection of starter is performed by a fuzzy controller SELECT-STARTER embedded in the SDN controller. Its design is based on the following requirements.

4.4 | Fewer burdens should be imposed on nongeocast nodes to transmit geocast message to starter

Current zone should not consume much energy in transferring the geocast message from geocast source/peripheral to the starter. This energy is just the cost of setting foundation of actual geocast operation. Best case happens when starter is at minimum possible energy distance from geocast source/peripheral. A node n_j is said to be at minimum possible energy

```
Function  Geocast-member-chk (node nᵢ, Geocast-zone AZ)
 Begin
   If (AZ. shape = circle) then
    Begin
       dec = (xᵢ(t) – h)² + (yᵢ(t) – k)² – r²
         /* (h,k) is center of AZ and its radius of r */
       If (dec < = 0)
         Insert ( AZ. Member, nᵢ)
   End
    If (AZ. shape = ellipse) then
     Begin
       dec = {(xᵢ(t) – h)/a}² + {(yᵢ(t) – k)/b}² - 1
         /* (h,k) is center of AZ; length of major and minor axis are a and b, in that order */
       If (dec < = 0)
         Insert ( AZ. Member, nᵢ)
    End

     If (AZ. shape = polygon) then
       Begin
         num = number of edges of the polygon AZ
         edge = array of edges of polygon AZ
         int k = 0

          for ( i=1; i<=num; i++)
            Begin
               If(edge[i].miny <= yᵢ(t) and edge[i].maxy <= yᵢ(t)) then
                Begin
                /* edge[i].miny and edge[i].maxy specify minimum and maximum y coordinate values
                corresponding to polygon edge edge[i]. x coordinates associated to edge[i].miny and
                edge[i].maxy are edge[i].minyx and edge[i].maxyx, in that order  */
                 Point intr
                 int slp, found = 0
                   slp = (edge[i].minyx - edge[i].maxyx)/( edge[i].miny - edge[i].maxy)
                  /* slp is slope of the polygonal edge edge[i] */
                   Intr[k].x = edge[i].minyx + (yᵢ(t) - edge[i].miny) × slp
                   Intr[k].y = yᵢ(t)
                  /* Intr[k] is intersection point of the line y = yᵢ(t) with the edge[i] */
                   k = k + 1
                   found = search (intr-array, Intr[k-1])
                    /* the above-mentioned search function returns 1 if last computed intersection point is found
                      Otherwise 0 is returned */
                    if (found = 0) then
                    insert (intr-array, Intr[k-1])
                End
               Int lim = k
               Ascend-bubble (intr-array, lim)
               /* Ascend-bubble applies bubblesort algorithm on intr-array to arrange intersection points in as
            of x coordinates */
                for (i=1; i<=lim; i=i+2)
                  Begin
                     if (xᵢ(t) >= intr-array[i].x and xᵢ(t) <= Intr[i+1].x) then
                    /* nᵢ is inside */
                      Insert ( AZ. Member, nᵢ)
                 End
         End Function
```

FIGURE 7 Pseudocode of determining whether a node is inside geocast region or not

distance edist$_t$(r,j) from one of its uplink neighbors n_r at time t, if Cartesian distance dist$_t$(r,j) between those nodes at time t, is equal to minimum possible physical distance min-dist between any two nodes in the network and minimum receive power mrp(j) is minimum.

Considering all nodes in the network. These conditions are formally written in Equations (9) and (10) whereas value of minimum possible energy distance is computed in (11).

$$\text{dist}_t(r,j) = \min-\text{dist}, \tag{9}$$

$$\text{mrp}(j) = \min \forall \text{mrp}(y), \text{s.t.,}\ \ y = 1 \text{ to } |Z|, \tag{10}$$

$$\text{edist}_t(r,j) = \{\text{mrp}(p)\ \ \min-\text{dist}^2\}/\text{C1}, \tag{11}$$

where Z is the set of all nodes in the current zone.

Extensive forwarding load is imposed on nongeocast members in propagating geocast message to starter, if all nongeocast members of current zone are involved in the task. (Z-A) is the set of nongeocast members in current zone. If distance between each pair of consecutive non-geocast members in the current zone is equal to maximum possible radio-range $R_{\max}$ in the network and minimum receive power of all of them is equal to max-mrp or maximum of minimum receive power considering all nodes in the network, then upper limit of energy consumption $\mathrm{upc}_{a,b}(t)$ of the route from geocast source/peripheral n_a to starter n_b, is given by (12).

$$\mathrm{upc}_{a,b}(t) = |Z - A| \max - \mathrm{mrp}R^2{}_{\max}/\mathrm{C1}, \tag{12}$$

where max - mrp = $\forall$ mrp(y), s. t. , $y = 1$ to $|N|$.

N is the set of all nodes in the network, including all zones.

4.5 | Starter should have good connectivity inside the geocast region

If a huge number of geocast members are reachable from a starter, then number of starters will be reduced and with that, forwarding load on nongeocast members will also be reduced. This requires good reachability for starters and also it will be delay efficient if nodes are placed at high levels in minimum energy transmission tree. Gain in terms of energy saving is also possible if certain conditions are satisfied.

Nodes should be at small hop distances from the starter so that depth of the tree is decreased. For example, consider the topologies of Figure 8A,B. Node n_i is starter in both the topologies of these two figures. It receives unicast message from geocast source or appropriate peripheral at time t. its only successor n_{i+1} receives the same at timestamp $(t + \Delta t1)$; similarly, n_{i+2} receives that geocast message at time $(t + \Delta t1 + \Delta t2)$; and so on upto n_{i+j} which receives geocast message at time $(t + \Delta t1 + \Delta t2 + \ldots + \Delta tp)$. In general, it can be said that timestamp tm_{i+j} $(2 \leq j \leq p)$ at which n_{i+j} receives the broadcast message, is given by (13),

$$\mathrm{tm}_{i+j} = \begin{cases} t + \Sigma\Delta t_j & \text{s.t. } j = 1 \text{ to } p \text{ and if } j > 0 \\ t & \text{if } j = 0 \end{cases}. \tag{13}$$

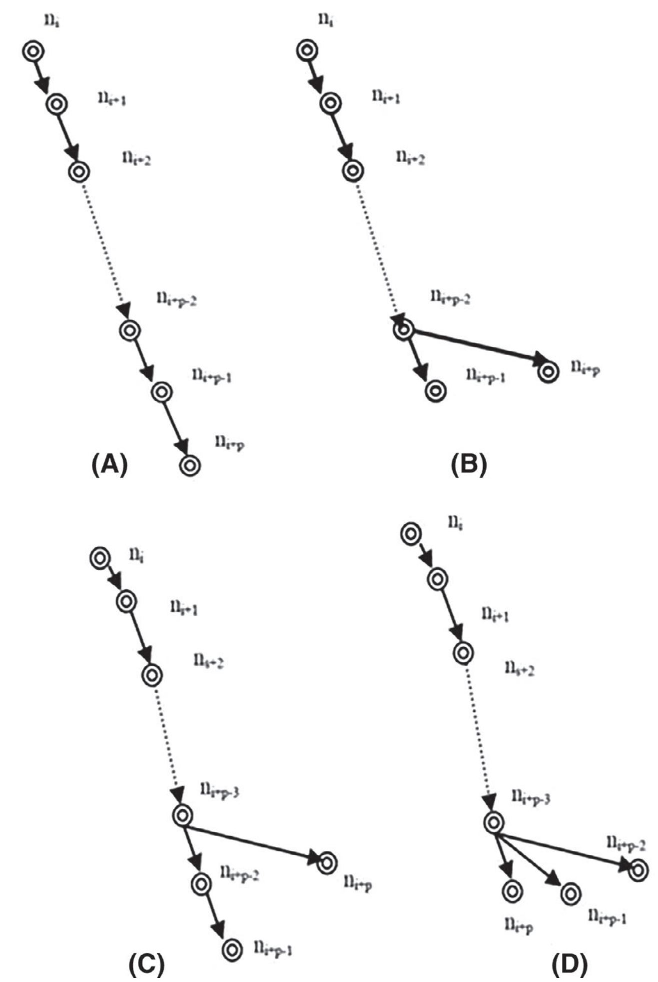

FIGURE 8 (A) Original topology. (B) Elevated topology. (C) Illustration of cascading elevation. (D) n_{i+p} is two levels up

As far as elevation of nodes is concerned, n_i and n_{i+1} cannot move up in the transmission tree because n_i is the starter itself; it is root of the transmission tree and no level is possible before the root. Similarly, n_{i+1} is an geocast member blessed with the best possible parent, that is, the starter. Therefore, only nodes from n_{i+2} to n_{i+p} can try to get attached to better father. Let a node n_{i+z} $(3 \le z \le p)$ moves up one level and gets attached to n_{i+z-2} while its previous predecessor was n_{i+z-1}. But one question arises here, that is, whether the children (if any, both single and multi-hop) of n_{i+z} will also move along with it so that predecessor of n_{i+z+1} remains n_{i+z} or n_{i+z-1} will become new predecessor of n_{i+z+1}. Movement of multiple nodes will incur significant kinetic energy cost. So, it is better to move leaf nodes (nodes without children), especially those who are lonely, that is, without siblings. That will lead to substantial energy savings because after elevation of the lonely leaf node, its predecessor will be completely relieved from forwarding geocast message. Hence, in this article only lonely leaf nodes are elevated. Cascading elevations are also possible in this kind of network structure when raising up a node triggers raising of its immediate predecessor. Particularly if a lonely leaf geocast member climbs up the transmission tree (at least two steps) then its immediate predecessor becomes childless and it can also move up. This is illustrated in Figure 7C,D.

Proposed geocast algorithm encourages elevation of only lonely leaf nodes. Let, n_{i+p} in Figure 8A detaches itself from its present uplink neighbor n_{i+p-1} and links with n_{i+p-2}.

Resultant topology appears in Figure 7B. Assume that, new and old timestamps at which n_{i+p} receives the geocast message are tm-new$_{i+p}$ and tm-old$_{i+p}$. Values of these are shown below in Equations (14) and (15).

$$\text{tm-new}_{i+p} = t + \Sigma\Delta t_j, \text{s.t.},\ j = 1 \text{ to } p - 1. \tag{14}$$

$$\text{tm-old}_{i+p} = t + \Sigma\Delta t_j, \text{s.t.},\ j = 1 \text{ to } p - 1 \tag{15}$$

Therefore improvement in delay is given by (tm-new$_{i+p}$- tm-old$_{i+p}$), that is, Δt_j.

Corresponding gain in energy is the difference between energy cost in old position and energy cost in new position. When n_{i+p} was the successor of n_{i+p-1}, transmission energy required by n_{i+p-1} is given by edist$_t(i+p\text{-}1,i+p)$ and defined in (16). After n_{i+p} got attached to n_{i+p-2}, transmission energy required by n_{i+p-2}, is defined in (17). But in order to change desired predecessor, n_{i+p} had to invest kinetic energy too. If old and new positions of n_{i+p} are (x-old$_{i+p}$, y-old$_{i+p}$) and (x-new$_{i+p}$, y-new$_{i+p}$), respectively, then velocity of the node was vl$_{i+p}$ appears in (18).

This is based on the assumption that transition took place in time duration te$_{i+p}$ and throughout the time it moved with uniform velocity. Corresponding kinetic energy ken$_{i+p}$ is shown in (19). mass$(i+p)$ is mass of n_{i+p}.

In Figure 8C, n_{i+p} has moved two levels up, that is, its predecessor is n_{i+p-3}. Now n_{i+p-1} has become a lonely leaf and it can move up now. This is cascading elevation. Figure 8D shows the situation where n_{i+p-1} has moved one level up and has become successor of n_{i+p-3}. Now n_{i+p-1} does not have any lonely leaf child.

Input parameters of SELECT-STARTER are dist-entr, connect-eff, res-eng-quotient. These are described below in (16) to (19).

$$\text{edist}_t(i+p-1, i+p) = \{\text{mrp}(i+p)\ \text{dist}_t{}^2(i+p-1, i+p)\}/\text{C1}. \tag{16}$$

$$\text{edist}_t(i+p-2, i+p) = \{\text{mrp}(i+p)\ \text{dist}_t{}^2(i+p-2, i+p)\}/\text{C1}. \tag{17}$$

$$\text{vl}_{i+p} = \text{pos}_{i+p}/\text{te}_{i+p}. \tag{18}$$

$$\text{s.t. } pos_{i+p} = \sqrt{\{(x - old_{i+p} - x - new_{i+p})^2 + (y - old_{i+p} - y - new_{i+p})^2\}}$$

$$\text{ken}_{i+p} = 0.5\ \text{mass}(i+p)\ \text{vl}^2{}_{i+p}. \tag{19}$$

Overall energy consumption in the network will reduce if the following condition mentioned in (20) is true.

$$\text{edist}_t(i+p-1, i+p) > (\text{edist}_t(i+p-2, i+p) + \text{ken}_{i+p}), \tag{20}$$

dist-entr(i,t) is the ratio of overall energy consumption in the least energy consuming path that does not include any geocast router, from geocast source or appropriate peripheral n_s to an geocast member n_i, with respect to maximum of the same from n_s to all geocast members in the current zone, at time t. Mathematically this is expressed in (21).

$$\text{dist-entr}(i,t) = \text{eng-dist}_{s,i}(t)/\{\max \forall \text{eng-dist}_{s,i}(t)\}, n_i \in A, \tag{21}$$

eng-dist$_{s,i}(t)$ is the amount of energy consumed in minimum energy path excluding all geocast member routers from n_s to n_i at time t. It is formulated in (22) based on the assumption that minimum energy path from n_s to n_i at time t is route$_{s,i}(t)$ such that,

$$\text{route}_{s,i}(t)\text{: } n_s = n_u \rightarrow n_{u+1} \rightarrow n_{u+2} \ldots n_{u+k(i)-1} \rightarrow n_{u+k(i)} = n_i,$$

where all nodes from n_{u+1} to $n_{u+k\text{-}1}$ are routers that do not belong to the geocast region. Number of nodes in route$_{s,i}(t)$ is $(k(i)$+1).Then,

$$\text{eng-dist}_{s,i}(t) = \Sigma \text{edist}(u + k' - 1, u + k'), \text{s.t.}, k' = 1 \text{ to } k(i). \tag{22}$$

From the formulation in (22), it is evident that dist-entr(i,t) lies between 0 and 1. Values close to 0 indicate that transferring geocast message from geocast source or appropriate peripheral node n_s to starter candidate n_i, does not burden nongeocast members much. Hence, from this perspective, n_i is in a favorable situation to become a router. Connect-eff(i,t) is a parameter that specifies topological efficiency of the not-yet-constructed transmission tree rooted at n_i at time t. It is high if a good number of geocast members are reachable from n_i without getting down much deep. That is, distance between n_i and farthest geocast member reachable from it, should be small. It is formulated in (23).

$$\text{connect-eff}(i,t) = \text{f-depth}(i,t)\ \text{f-reach}(i,t), \tag{23}$$

where $\text{f-depth}(i, t) = \sqrt{}(1 - \text{max-dep}_i(t)/(|A| - 1))$, and $\text{f-reach}(i, t) = \text{max-rch}_i(t)/|A|$.

max-dep$_i(t)$ is maximum possible depth in transmission tree rooted at n_i at time t. Maximum value of max-dep$_i(t)$ is $(|A| - 1)$ where all geocast members will act as routers to carry the message from n_i to the geocast member at maximum depth. In this case, all nodes in the transmission tree will have only one child. max-rch$_i(t)$ is the total number of nodes reachable from n_i at time t.

The relation reach-able of reachability is reflexive and transitive, but not symmetric. A node is always reachable from itself; that is, $\forall t$, $n_i \in$ reach-able(n_i, t). Similarly, if $n_i \in$ reachable(n_i, t) and $n_k \in$ reachable(n_j, t), then $n_k \in$ reach-able(n_i, t) in two hops, that is, the relation is transitive. But $n_i \in$ reachable(n_i, t) does not necessarily mean the reverse that is, $n_i \in$ reachable (n_i, t) because all links in a zone are not bidirectional. At most $|A|$ number of nodes can be reachable from a starter.

From the formulation in (23) it is evident that connect-eff(i,t) lies between 0 and 1. Values close to 1 denote small highest depth in the transmission tree with a good number of reachable nodes. Pseudocode to compute connect-eff is shown in Figure 9 which is continued to Figure 10.

res-eng-quotient(i,t) specifies residual battery energy of n_i at time t compared to its maximum energy. It is expressed in (24).

$$\text{res-eng-quotient}(i,t) = \text{r-en}_i(t)/\text{m-en}_i. \tag{24}$$

Values of this parameter close to 1 indicate high residual energy that is expected to yield a stable root for the transmission tree.

Entire range (0-1) of all of these parameters, is uniformly divided into four ranges (0-0.25 is denoted by a, 0.25-0.50 is denoted by b, 0.50-0.75 is c whereas the next higher range 0.75-1.00 is denoted by d). The fuzzy is shown in Tables 1 and 2. In Table 1 it can be seen that dist-entr = *a or b*, then it means that nongeocast nodes need not spend much battery power to carry geocast message to potential starter, from the geocast source or appropriate peripheral. This is good for profile of the starter candidate. On the other hand, connect-eff should be.

Rule bases of SELECT-STARTER are as under.either c or d for a candidate competing to become a starter. It indicates good reachability without traversing downmuch. Hence (a,d), that is, ath row and dth column of Table 1, is best possible combination for the couple dist-entr and connect-eff. This combination produces best value for $t1$, that is, $t1 = d$. As dist-entr starts to increase and becomes b, then for same connect-eff, $t1$ reduces to c. Similarly, as dist-entr becomes c without changing connect-eff, $t1$ will reduce to c, and so on. This means that for same reachability and depth, performance of a starter becomes better with reduction is dist-entr. Also for a given dist-entr, increase in connect-eff lead to better performance of a starter. For example, consider the columns of Table 1. First column is (a, b, c, d). So, value of $t1$ is

```
Function Compute-reachability ( Graph graph, int starter)
 Begin
  Queue q, q′, q″
  /* At the beginning, it is required to create the list of downlink and uplink neighbors of all nodes in geocast region */
  int tot-reach = 0
 /* tot-reach is total number of nodes reachable from starter */
 for each node np∈A
  begin
   for each node nv∈A
    begin
     if distt(v,p) < radv then
     begin
      /* The condition distt(v,p) < radv specifies np is within radio-range of nv. So, np is a downlink neighbour of nv
      And nv is an uplink neighbour of  np*/
      Insert (v → dneighbourlist, p)
      Insert (p → uneighbourlist, v)
  End
 End
 End
 for each node np ∈ A
  begin
   copy ( p → uneighbourlist, p → predecessorlist)
  end
  insert(q, starter)
    /* identification number of starting node is inserted in queue q */
  int curlevel = 1
  int node-at-level[] = malloc(sizeof(int) × num-of-geocast-mem)
   /* num-of-geocast-mem is |A| */
  /* node-at-level[1] is number of successors of starter at level 1, similarly, node-at-level[2] is number of successors of starte
  level 2 and so on */
     While (!isEmpty(q) and   !isEmpty(q′))
       Begin
        /* isEmpty is a function that checks whether a queue is empty or not */
          /* If the condition (!isEmpty(q) and   !isEmpty(q′)) is true then it means that there are more nodes to visit */
         If (isEmpty(q) and   !isEmpty(q′))  then
          Begin
            /* Successors of current node are waiting in queue q′. These elements need to be copied in q for further processing
            copy (q, q′)
            insert(q″, curlevel)
            merge(q″,q′)
            insert(q″, -1)
             /* both q and q′ are containing next level nodes to traverse in current pass. All nodes are stored in q″ with their
before them. -1 is the delimiter after each level number */
            curlevel = curlevel + 1
            node-at-level[curlevel] = count[q′]
            delete(q′);
         End
          currentnode = dequeue(q)
          /* currentnode is the first one in queue q */
         currentnode→visited = TRUE
         temp = currentnode → dneighbourlist
          while (temp)
            begin
              /* there are more neighbours of the current node */
             nde = temp → node
              /* nde is current neighbor of the node currentnode */
```

FIGURE 9 First part of pseudo code of the function to compute reachability

increasing from top to bottom, that is, with increase in connect-eff; similarly second (a, b, b, c), third (a, a, b, b), and fourth (a, a, a, b) columns also yield the similar observation.

Ultimate output starter-merit increases with increase in $t1$ as well as res-eng-quotient, as can be seen in Table 2. For example, observing the flow along the rows, four rows (a, a, a, b), (a, b, b, b), (b, b, c, c) and (c, c, d, d) can be seen; that is, starter-merit becomes better with increase in value of $t1$. Similarly, considering the columns, it can be seen that starter-merit improves with improvement in res-eng-quotient.

First, second, third, and fourth columns are (a, a, b, c), (a, b, b, c), (a, b, c, d) and $(b. b, c, d)$, that is, improvement is noticed while getting down in each column. Getting down means increase in res-eng-quotient.

After the starter is decided, a minimum energy transmission tree is formed from the starter to each of its nonstarter reachable geocast members. Let $\lambda(i, j)$ be the set of routes from starter n_i to a node n_j which is reachable from n_i. Then a route $\text{ROUTE}_w(i, j)$ *s.t.* $1 \le w \le |\lambda(i, j)|$ will be chosen for communication from n_i to n_j, if condition in (25) is true. This means that n_j is under supervision of n_i till now.

$$\begin{aligned}&\text{eng-ct}(\text{ROUTE}_w(i,j)) \le \forall \text{eng-ct}(\text{ROUTE}_{w'}(i,j)),\\&\text{s.t.}, w' = 1 \text{ to } \mid \lambda(i,j) \mid .\end{aligned} \tag{25}$$

```
if ( nde → visited == FALSE)
 Begin
   /* nde is not yet visited */
   insert (nde → predecessorlist, currentnode)
   merge (nde → predecessorlist, currentnode → predecessorlist)
    /* merge (L1, L2) is a function that concatenates L1 and L2 and stores the result in L1 */
   /* Both insert and merge functions take care of duplicates. currentnode will be inserted in nde → predecessorlist
      Only if it not already there */
   /* Similarly each element of currentnode → predecessorlist is searched in nde → predecessorlist. If that already
     exists in nde → predecessorlist, then it will not be inserted in currentnode → predecessorlist. Otherwise insertion will be done
    */
     enqueue ( q′, nde)
     /* nde is a node in the next level and so it is in queue q */
     temp = temp → next
    end
   end
  end
 for (lim =1; lim<=curlevel; lim++)
  begin
   tot-reach = tot-reach + node-at-level[lim]
   end for
 end function
```

FIGURE 10 Second part of pseudo code of the function to compute reachability

TABLE 1 Fuzzy combination of dist-entr and connect-eff producing temporary output $t1$

dist-entr→connect-eff↓	*a*	*b*	*c*	*d*
a	*a*	*a*	*a*	*a*
b	*b*	*b*	*a*	*a*
c	*c*	*b*	*b*	*a*
d	*d*	*c*	*b*	*b*

TABLE 2 Fuzzy combination of $t1$ and res-eng-quotient producing ultimate output starter-merit

t1→res-eng-quotient↓	*a*	*b*	*C*	*d*
a	*a*	*a*	*A*	*b*
b	*a*	*b*	*B*	*b*
c	*b*	*b*	*C*	*c*
d	*c*	*c*	*D*	*d*

If multiple routes suffer from same energy consumption, then any one of them is selected randomly. This means that there will be only one route from the starter to each nonstarter geocast member.

4.6 | Elevating lonely leaves

For elevating lonely leaves, green geocast allocates a limited time τ-reloc. Suppose n_e is a lonely leaf having highest velocity capacity vel-max$_e$. Then among all of its predecessors (available from predecessor list computed in Pseudocode of Figures 9 and 10, a node n_j will be capable of giving shelter to n_e if closest point on radio-circle of n_j can be reached by n_e within time τ-reloc. This condition appears in (26).

$$(\text{dist}_t(i, e) - \text{rad}_j) \leq (\text{vel-max}_e \times \tau - \text{reloc}). \tag{26}$$

Maximum possible energy gain will be the difference of transmission energy gain and kinetic energy invested. Both are expressed in Joules.

Assume that at current time t, n_e is under the supervision of n_i with their present distance being $\text{dist}_{i,e}(t)$. Then energy consumption of n_i to send an unit length message packet to n_e is given by (27).

$$\text{edist}_t(i,e) = \{\text{mrp}(e)\ \text{dist}_t{}^2(i,e)\}/\text{C1}. \tag{27}$$

Minimum possible distance between any uplink neighbor of n_e and n_e, is unit distance or 1. Therefore, if n_e becomes maximally close to an uplink neighbor n_j, then $\text{edist}_t(j,e)$ will be formulated as (28).

$$\text{edist}_t(j,e) = \{\text{mrp}(e)/\text{C1}\}. \tag{28}$$

So, maximum possible transmission energy gain MAX-TRANS-ENG in the network (please note that required transmission energy of n_i will decrease and for n_j, it will increase but if reduction in transmission energy of n_i is higher than increase in the same for n_j along with kinetic energy of n_e, then relocation of n_e will be beneficial for the network) is given by (29).

$$\text{MAX-TRANS-ENG} = \text{mrp}(e)(\text{dist}_t^2(i,e) - 1)/\text{C1}. \tag{29}$$

So, during elevation of n_e, otential supervisor (or predecessor) of n_e is chosen with closest radio-circle entry point or base point. Underlying intention is to reduce kinetic energy consumption of n_e.

The concept is illustrated in Figure 11. n_e moves to potential supervisor n_j in a straight line and the line intersects radio-circle of n_j at point (h,k). Hence, (h,k) is called entry point of n_e to radio-circle of n_j. Another name of this point is base point. Coordinates of base point can be easily calculated with law of quGDRatic from the followings:

$$(h - x(t))^2 + (k - y_j(t))^2 = \text{rad}_j{}^2, k = y_e(t) + m'(h - x_e(t))$$

where $m = (y_e(t) - y_j(t))/(x_e(t) - x_j(t))$ Kinetic energy $\text{ki-eng}_{h,k}(e,t)$ required by n_e at time t to arrive at base point of n_j, is shown in (30).

$$\text{ki-eng}_{h',k'}(e,t) = \{\text{mass}(e) \times \text{vel}^2{}_{i,j}(e,t)\}/2, \tag{30}$$

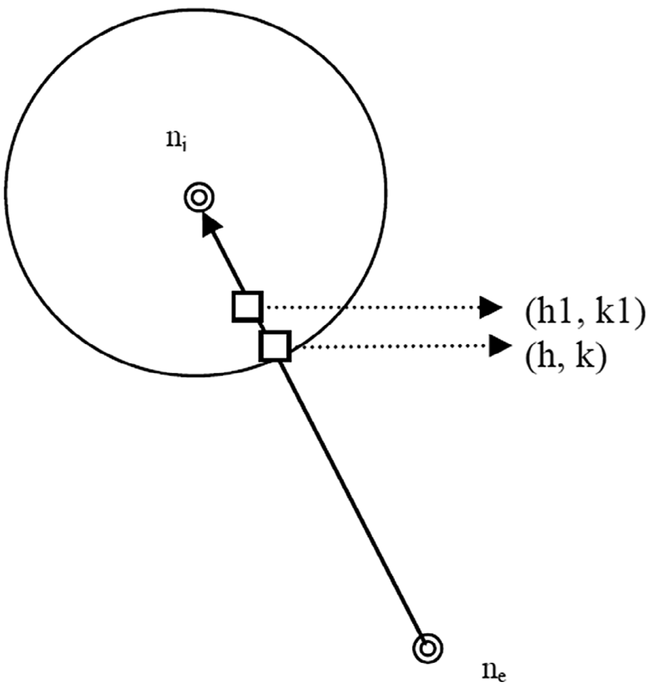

FIGURE 11 n_e gets elevated to radio-circle of n_j

where mass (e) is the mass of n_e and $\text{vel}_{i,j}(e,t)$ is the required velocity of n_e to reach base point of radio-circle of n_j from n_i in time duration τ-reloc. So, as in (31)

$$\text{vel}_{i,j}(e,t) = D_{e,t}(h,k)/\tau\text{-reloc}, \tag{31}$$

such that $D_{e,t}(h,k) = \sqrt{\{(x_e(t)-h)^2 + (y_e(t)-k)^2\}}$.

Definitely, $\text{vel}_{i,j}(e,t) \le \text{vel-max}_e$.

Transmission energy $\text{tr-eng}_{h,k}(e,t)$ in this context is given by (32),

$$\text{tr-eng}_{h,k}(e,t) = \text{mrp}(e)\ (\text{dist}_t^2(i,e) - D^2{}_{j,t}(h,k))/\text{C1}. \tag{32}$$

Overall energy gain $\text{ov-eng}_{h,k}(e,t)$ as a result of relocation or elevation of n_e, is expressed in (33).

$$\text{ov-eng}_{h,k}(e,t) = \text{tr-eng}_{h,k}(e,t) - \text{ki-eng}_{h,k}(e,t). \tag{33}$$

Further elevations of n_e are possible if it generates more energy gain. For example, let $(h1, k1)$ be the point that is next to (h, k) toward n_j. Assume that $(h1, k1)$ divides the line from (h,k) to $(x_j(t), y_j(t))$ in the ratio $(\text{rad}_j\text{-}\alpha)$: α. Then,

$$h1 = (x_j(t)\alpha + (\text{rad}_j - \alpha)h)/\text{rad}_j.$$

$$k1 = (y_j(\text{t})\alpha + (\text{rad}_j - \alpha)k)/\text{rad}_j$$

Elevation to $(h1, k1)$ is possible if the following two conditions are true:

(i) $D_{e,t}(h1, k1) \le (\text{vel-max}_e \times \tau\text{-reloc})$
(ii) $\text{ov-eng}_{h,k}(e,t) \le \text{ov-eng}_{h1,k1}(e,t)$.

If moving toward n_j continues to produce more overall energy gain, then ultimately elevation has to stop when n_e arrives at unit distance from n_j. Overall energy gain at last step of elevation will be termed as optimum overall energy gain.In case of multiple such predecessors producing similar optimum overall energy gain, the one that yields least interference to the network after inclusion of n_e, is chosen as optimum.

Procedure to compute interference is shown below with the help of Figure 12. In Figure 12, it can be seen that n_e lies within the neighborhood of both n_i and n_j. This structure is quite possible in case of energy aware transmission tree. Although in this tree only one uplink neighbor is responsible for sending geocast message to a node but chances of interference exist. For example, say n_i is responsible for sending geocast message to n_k and n_j is responsible for sending geocast message to n_e, but still n_e may overhear, especially when n_i has a downlink neighbor n_k close to n_e, resulting in collision. Multiple such interfering may happen. The intention is to find out a general expression of such interference.

Maximum and minimum possible transmission powers in the geocast region are denoted as max-trans-power and min-trans-power and defined in (34) and (35).

$$\text{max-trans-power} = \{\max(\forall \text{mrp}(i))\}\ R^2{}_{\max}\ \text{ s.t.}, n_i \in A. \tag{34}$$

$$\text{min-trans-power} = \{\min\ (\forall \text{mrp}(i))\}, \text{s.t.}, n_i \in A. \tag{35}$$

Average transmission power P_{avg} of any node in the geocast region is given by (36).

$$P_{\text{avg}} = (\text{max-trans-power} + \text{min-trans-power})/2. \tag{36}$$

To compute composite interference after inclusion of n_e in radio-circle of n_j, the authors consider a differential area $(\beta\ d\beta\ d\phi)$ where n_e is at distance β from n_j and the straight line connecting n_j and n_e makes an angle ϕ with the straight line connecting n_i and n_j. Assuming uniform distribution of nodes in the geocast region, amount of interference faced by n_e as receiver, is denoted as dI and defined in Equations (37) and (38).

$$dI = (P_{\text{avg}}/\beta^2)f'|A|\beta d\beta d\phi\ , \tag{37}$$

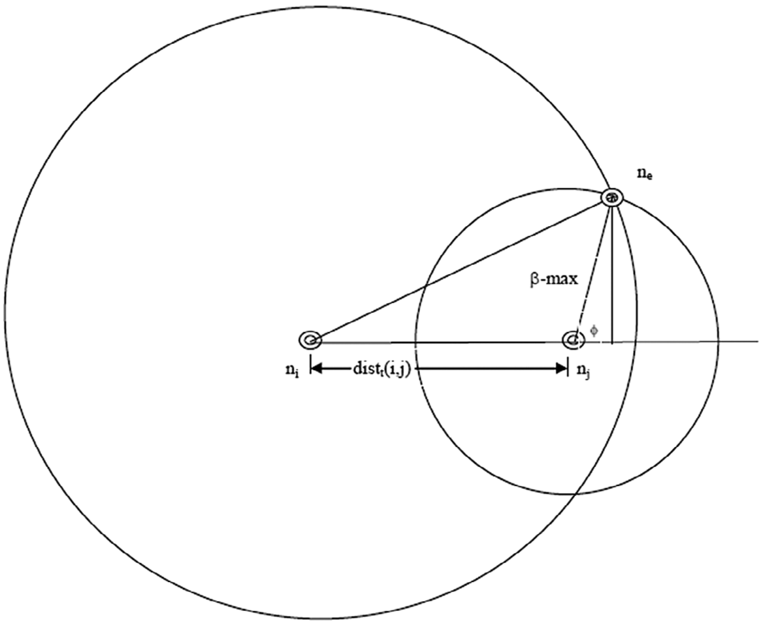

FIGURE 12 Illustration of interference

$$dI = (P_{\text{avg}}/\beta) f' |A| \, d\beta d\phi, \tag{38}$$

where f' is fraction of senders in the geocast region. So, f' $|A|$ is number of senders in the geocast region. Please note that $0 \le f' \le 1$. Therefore, expected interference $E(I)$ is formulated in (39).

$$E(I) = \iint_{0}^{2\pi \beta-\max} (P_{\text{avg}}/\beta) f' |A| \, d\beta d\phi, \tag{39}$$

where minimum value of β is 1. β will be at its maximum if n_e lies on the periphery of radio-circle of n_i, because, after all, n_e should remain a downlink neighbor of n_i so that n_i can create an interference. P_{avg}, f' and |A| are constants.

$$\text{So}, E(I) = 2P_{avg} f' |A| \int_0^{\pi} \log(\beta - \max) \, d\phi. \tag{40}$$

Value of β-max needs to be put in (40) to compute $E(I)$. It is done by drawing a perpendicular from n_e to the straight line connecting n_i and n_j. This perpendicular touch the straight line from n_i to n_j at point Y. Noted that distance from n_i to n_e is rad_i. Hence, from Figure 12 it may be seen that,

$(\text{dist}_t(i,j) + \beta - \max\cos\phi)^2 + (\beta - \max\sin\phi)^2 = r$, that is, $\beta - \max^2 + 2\beta - \max\cos\phi\text{dist}_t(i,j) + \text{dist}^2{}_t(i,j) - \text{rad}_i{}^2 = 0$,

So, $\beta - \max = -\text{dist}_t(i,j)\cos\phi \pm \sqrt{}(\text{dist}^2{}_t(i,j)\cos^2\phi + \text{rad}_i{}^2)$.

β-max cannot be negative because it is a distance. Therefore,

$$\beta\text{-}\max = -\text{dist}_t(i,j)\cos\phi + \sqrt{}(\text{dist}^2{}_t(i,j)\cos^2\phi + \text{rad}_i{}^2). \tag{41}$$

To calculate $E(I)$, the value of β-max in (41) is put in (40).

4.7 | No cycles can exist in energy efficient transmission forest

After selecting a suitable starter and elevating all the lonely leaves, exactly one minimum energy consuming route is constructed from selected starter to each nonstarter downlink neighbor. In case of more than one minimum energy paths, any one is randomly selected for communication. So, at this point of time, each node is under supervision of exactly one uplink neighbor.

Please note that, no cycle can be generated in this process. This is proved using the Lemmas 1 and 2.

Lemma 1. *There can be no cycle in energy efficient transmission tree involving the starter.*

Proof. Consider Figure 13 where n_i is the starter and its downlink neighbor is n_j. Downlink neighbors of n_j are n_v and n_u. n_k is a downlink neighbor of n_u. The cycle is $n_i \rightarrow n_j \rightarrow n_u \rightarrow n_k \rightarrow n_i$. Therefore n_k is an geocast member which is uplink neighbor of the starter which is not possible. So, there cannot be any geocast member which is uplink neighbor of the starter. ∎

Lemma 2. *There can be no cycle in energy efficient transmission tree involving nonstarter geocast members.*

Proof. Consider Figure 14 where n_i is the starter and its downlink neighbors are n_j and n_k. Without any loss of generality it has been assumed that there is a cycle $n_k \rightarrow n_p \rightarrow n_q \rightarrow n_r \rightarrow n_k$. Please note the two paths $n_i \rightarrow n_k \rightarrow n_p \rightarrow n_q \rightarrow n_r \rightarrow n_k$, $n_i \rightarrow n_k$, in energy efficient transmission tree in Figure 14. So this means that minimum energy consuming paths from n_i to n_k are $n_i \rightarrow n_k \rightarrow n_p \rightarrow n_q \rightarrow n_r \rightarrow n_k$ and $n_i \rightarrow n_k$. According to the logic of creating energy efficient transmission tree, if more than one minimum energy paths are there then only one of them is selected randomly and inserted in energy efficient transmission tree but not both of them. But in Figure 14, both the paths $n_i \rightarrow n_k \rightarrow n_p \rightarrow n_q \rightarrow n_r \rightarrow n_k$ and $n_i \rightarrow n_k$ are appearing which contradicts the principle of forming energy efficient transmission tree. Therefore, combining the findings in Lemmas 1 and 2, it can be concluded that no cycles can exist in an energy efficient transmission tree.

After constructing an energy efficient transmission tree, it needs to be checked whether all geocast members are included in it. If some members are left, then again a starter has to be selected among them and more transmissions

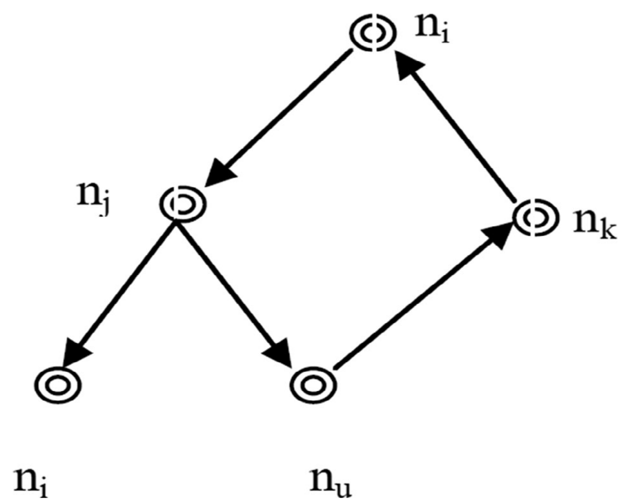

FIGURE 13 Cycle involving starter

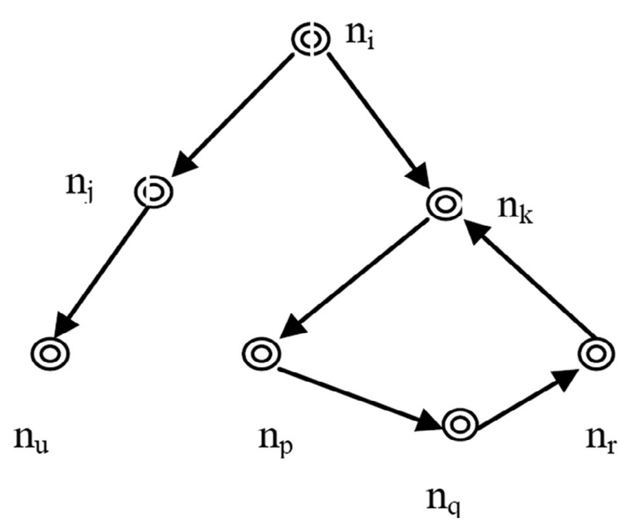

FIGURE 14 Cycle without starter

trees are constructed out of those geocast members which are not included in any energy efficient transmission trees constructed so far. Minimum-sized energy efficient transmission tree consists of only one node and in that case, the node will be starter itself and like all starters, it will receive geocast message from geocast source or incoming geocast peripheral. ∎

Lemma 3 proves that there cannot be any multitree cycle in energy efficient transmission forest.

Lemma 3. *No cycle in energy efficient transmission forest can involve multiple energy efficient transmission trees.*

Proof. In Figure 15, two starters n_i and n_j are there. There exists a cycle $n_p \rightarrow n_q \rightarrow n_r \rightarrow n_e \rightarrow n_f \rightarrow n_p$ that involves two energy efficient transmission trees originating at starters n_i and n_j. Nodes in first energy efficient transmission tree are, n_i, n_u, n_v, n_e, n_f while nodes in second energy efficient transmission tree are n_j, n_p, n_q, and n_r. Please note that, in Figure 3, first energy efficient transmission tree is marked in red whereas the second energy efficient transmission tree is marked in blue.

Since the cycle $n_p \rightarrow n_q \rightarrow n_r \rightarrow n_e \rightarrow n_f \rightarrow n_p$ is already created then n_p is reachable from n_f although n_f belongs to the first energy efficient transmission tree whereas n_p belong to the second energy efficient transmission tree. This is impossible because if n_p is reachable from n_f, then n_p is reachable from n_i which is starter of the first energy efficient transmission tree. But if n_p is reachable from n_i then n_p should belong to first energy efficient transmission tree which contradicts the basic structure of different transmission trees. Two different energy efficient transmission trees cannot have any node in common. Hence, it can be concluded that no cycle in energy efficient transmission forest can involve multiple energy efficient transmission trees. ∎

4.8 | Interzonal geocast message tranfer

After processing in current zone is complete or if no geocast members reside in the current zone, then current zone sends geocast message to all the peripheral nodes except the one from which geocast message has arrived. In case of existence of multiple paths from incoming peripheral node to other peripheral nodes, the path that consumes minimum energy is selected. But if the geocast source itself is in the current zone and no geocast member resides in the current zone then geocast message is sent from source to all neighbor zones through minimum energy consuming paths. But if more than one peripheral node corresponds to a neighbor zone, then a peripheral node is chosen depending upon output of a fuzzy controller named GEOCAST PROPAGATOR. Input parameters of this fuzzy controller are res-eng-quotient and zone-connect. Res-eng-quotient is already formulated and zone-connect is mathematically expressed in (42).

$$zone - \mathrm{connect}(i, \mathrm{zn}, t) = \mathrm{ng}(i, \mathrm{zn}, t)/\Sigma\, \mathrm{ng}(j, \mathrm{zn}, t); n_j \in (\mathrm{intrfc}(\mathrm{zone}(n_i), \mathrm{zn})). \tag{42}$$

ng(i,zn,t) specifies number of neighbors of peripheral node n_i in zone zn at time t. zone(n_i) is identifier of the zone to which the node belongs. Intrfc(zone(n_i),zn) is the set of all nodes belonging to zone of n_i having at least one downlink neighbor in zone zn.

zone-connect (i, zn, t) lies between 0 and 1. If its value is high then it means that n_i has good connectivity with zone zn compared to other peripheral nodes in the zone of n_i, that has at least one neighbor in zone zn.

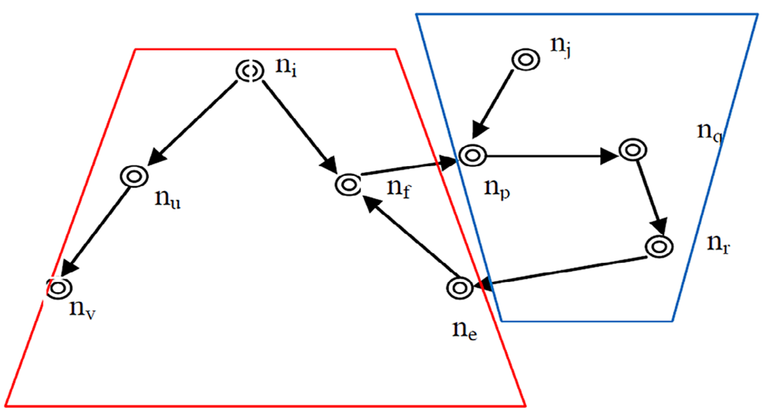

FIGURE 15 Cycle involving more than one energy efficient transmission tree

TABLE 3 Fuzzy combination of zone-connect and res-eng-quotient producing ultimate output peri-eff

zone-connect→res-eng-quotient↓	*a*	*B*	*c*	*D*
A	*a*	*A*	*b*	*B*
B	*a*	*B*	*b*	*C*
C	*b*	*B*	*c*	*D*
D	*b*	*C*	*d*	*D*

Fuzzy rule base combining the above two parameters is shown in Table 3. Output produced by Table 3 is called peri-eff which specifies efficiency of the corresponding peripheral node. Range division of zone-connect and peri-eff are same as res-eng-quotient. An efficient peripheral node is one that has high residual energy and good neighbor zone connectivity, corresponding to one particular neighbor zone. For a given zone-connect, if res-eng-quotient increases, then peri-eff also increases. Similarly, For a given res-eng-quotient, if zone-connect increases, then peri-eff also increases.

5 | SIMULATION EXPERIMENTS AND RESULTS

Simulation is performed in SDN framework and the communication protocol for WSN in each zone is IEEE 802.15.4. AMDOpteron processor is used with 16 GB hard disk. Used operating system is Linux version 2.6.32. Number of zones is 10. Total number of simulation runs is also 10. Number of nodes in each zone varies from 10 to 50. Total number of nodes in the network is 100, 150, 200, 250, 300, 350, 400, 450, and 500 in various simulation runs where network area is a square with length of each side being 500 m. Radio-range of each node varies from 10 to 50 m whereas initial energy ranges from 10 to 30 J. Size of each data packet is 512 bytes. Duration of each simulation run is 1000 seconds. Performance of our proposed green geocastis compared with GEAR,[39] EAGR,[47] GPSR,[16] EELIR,[18] and FRFZ,[26] which are popular geocast protocols in WSN only, because, as far as the authors know, no geocast protocol in SD-WSN framework has come out yet.

Simulation metrics are,

(i) Message Transmitted/Forwarded (MT/F): This is expressed as the sum of number of messages sent by all nodes in the network. Let msge(k) denotes total number of messages transmitted or forwarded by n_k throughout the simulation period. This does not have any unit.
$\text{MT/F} = \Sigma\ \text{msge}(k)\ n_k \in N$

(ii) Total energy consumption (TEC): This is the sum of energy consumed by all nodes in the network throughout the simulation period. Energy consumption is in millijoule.
$\text{TEC} = \Sigma(m\text{-}en_k\text{-}r\text{-}en_k).\ n_k \in N$

(iii) Geocast Delivery Ratio (GDR): This specifies percentage of geocast packets that were successfully delivered to geocast destinations to multiple geocast sessions. Let tr-pac be the number of packets transmitted or forwarded throughout the simulation. dlv-pac is the number of packets successfully delivered to geocast destinations throughout the simulation run. GDR cannot have any unit.

$$\text{GDR} = (\text{dlv} - \text{pac}/\text{tr} - \text{pac}) \times 100.$$

(iv) Average Geocast Delay (AGD): This is the time difference between origination of an geocast message and its successful delivery to the last geocast member. Let PAC be the set of all packets transmitted throughout the simulation. org(pack) and dlvr-last(pack) denote the timestamp of generation of the packet pack and timestamp of its delivery to the last geocast member, respectively.

$$\text{AGD} = \Sigma(\text{dlvr} - \text{last(pack)} - \text{org(pack)})/|\text{PAC}| \quad \text{pack} \in \text{PAC}$$

If one particular packet pack is not reachable at all, then (dlvr-last(pack) − org(pack)) is set to maximum delay in our dataset, considering all protocols competing here. Similarly, for each such packet, number of messages corresponding to the longest route from multicast source/ appropriate incoming peripheral (in case of green geocast) to an geocast member considering all competing protocols, is added to MT/F. Corresponding energy consumption is computed as (highest possible energy consumption per hop × number of hops in the longest route from multicast source/appropriate incoming peripheral (in case of green geocast) to a geocast member). This energy consumption is added to TEC calculated so far. Average delay is in ms.

(v) Percentage of Live Nodes (PLN): As the name suggests, this is percentage of live nodes in the network. Let N be set of all nodes in the network and among then, N1 be number of live nodes after some geocast operations. PLN does not have any unit. Then PLN is defined as,

$$PLN = (N1/N) \times 100.$$

5.1 | About competitors of green geocast

GEAR and GPSR follow perimeter flooding method which produce good packet delivery ratio only when the network is fairly populated. But under extremely high population geocast delivery ratio reduces due to high message contention and collision that generates because of redundancy corresponding to flooding in geocast region. EAGR is based on directional flooding. Hence in sparsely populated network, EAGR produces more GDR. But as network becomes extremely crowded, GDR produced by EAGR starts to drop more than others because of huge redundancy in geocast region (common to GEAR and GPSR) along with redundancy while traveling toward the geocast region (this redundancy is absent in GEAR and GPSR). Message cost and energy consumption in a sparse network is higher in GEAR and GPSR than EAGR and EELIR because although message cost corresponding to unreachable geocast members should be ideally infinity, but for ease of calculation and plotting it has been mapped to number of messages corresponding to the longest route from multicast source/appropriate incoming peripheral (in case of green geocast) to an geocast member, considering all protocols competing here. GEAR and GPSR suffer from a high number of unreachable nodes, so, message cost will be higher for them. EELIR builds energy efficient unicast routes to each geocast member. It avoids flooding during data packet transmission (as already mentioned, flooding is required only for route discovery) is completely avoided and therefore does not suffer from redundancy. But its message cost increases particularly when geocast source is far from geocast region. On an average, it can be seen that, when geocast region is not that populated, EELIR produces much lesser message cost than GEAR, EAGR and GPSR. The gap starts reducing when number of nodes in the network increase substantially.

5.2 | Green geocast vs GEAR

GEAR is a geocast protocol that chooses the downlink neighbor closest to centroid of geocast region. If more than one downlink neighbors are at minimum distance from the centroid of geocast region, compared to other downlink neighbors of the same node, then any one is randomly selected. There is no guarantee that the random selection is the most energy efficient one. After the geocast message reaches the first geocast member, it applies blind flooding inside the geocast region. For example, let n_i be the first geocast member at which geocast message has arrived. Then n_i forwards the geocast message to those of its downlink neighbors which are presently residing inside the geocast region. If all downlink neighbors of n_i are inside that region, then n_i can apply broadcast operation; otherwise, multiple unicasting has to be performed. This generates redundancy if one geocast member is within radio-circles of more than one geocast member. Such unnecessary retransmissions generating due to topological redundancies eat up significant amount of battery power in geocast members. Therefore, lifetime of geocast members decrease in GEAR. Also, there is no guarantee that all nodes in geocast region are reachable from n_i. If n_i has no downlink neighbor, then geocasting will stop immediately. Thus, geocast delivery ratio in GEAR is lesser than its competitors. Although GEAR applies flooding inside geocast region, but still that flooding is initiated by only one node and if its connectivity inside geocast region is poor, then percentage of successful delivery of geocast packet, will be low.

On the other hand in our proposed energy efficient geocast approach,the network is divided into multiple zones and each zone is controlled by one SDN controller. After receiving an geocast message, the incoming peripheral nodes informs corresponding SDN controller and the controller determines which nodes under his zone are inside the geocast region. Those nodes are geocast members in that zone. Since topological information of the zone is available to the controller, it can construct an energy efficient transmission forest. Flooding inside the geocast region is completely eliminated leading to huge saving in terms of message cost as well as energy in nodes. Energy efficient transmission forest consists of some energy efficient transmission trees where each nonstarter geocast member is under supervision of exactly one geocast member. Hence, redundancy is completely avoided, leading to great improvement in terms of messages transmitted or forwarded, as shown in Figure 16.

During selection of each starter, the criterion of reachability is given great importance; that is, the intention was to include as many geocast members as possible, in the geocast transmission tree being constructed. Hence, geocast delivery ratio is high in the proposed approach, as shown in Figure 18. Also, this reduces number of starters, especially when the network (or geocast region) in densely populated. For energy efficiency purpose, each node is connected to starter of the associated energy efficient transmission tree through minimum energy consuming paths.All these contribute to least energy consumption in green geocast, as shown in Figure 17. Lonely leaves are elevated wherever energy efficient uplink neighbor options are available. This reduces delay in geocast operation. This is shown in Figure 19.

5.3 | Green geocast vs FRFZ

FRFZ applies the technique of limited area flooding where the smallest rectangle circumscribing geocast region is targeted. Sender node broadcasts route-request toward this rectangle to find suitable nodes close to four such vertices. Then all those nodes broadcast the message toward the geocast region. This does not have any energy conserving technique inside geocast region, only limited area flooding is applied to find vertices of the circumscribing rectangle.

On the other hand, green geocast does not require flooding. Messages are delivered to geocast destinations through minimum energy paths leading to huge energy savings. More number of alive nodes guarantee almost hundred percent message delivery, as evident from Figures 16 to 21.

5.4 | Green geocast vs EAGR

EAGR is an geocast protocol where each router adjusts its transmission power so that it is just sufficient to reach the next hop router. Directional flooding is used for geocasting. In directional flooding each router targets all those vertices

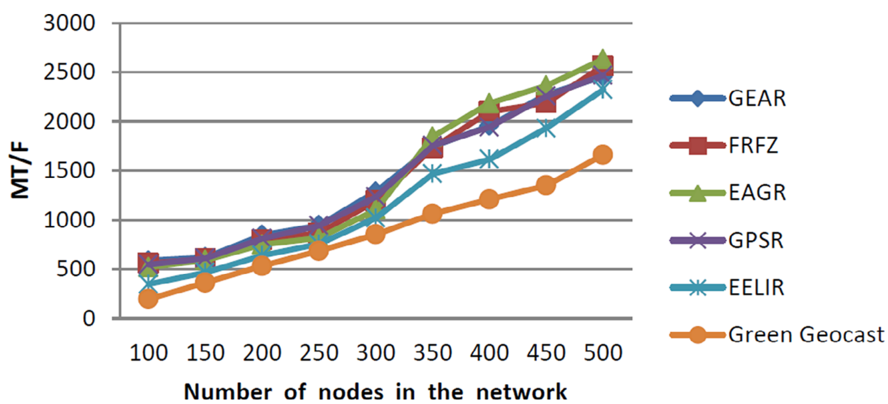

FIGURE 16 Message transmitted/forwarded vs number of nodes in the network

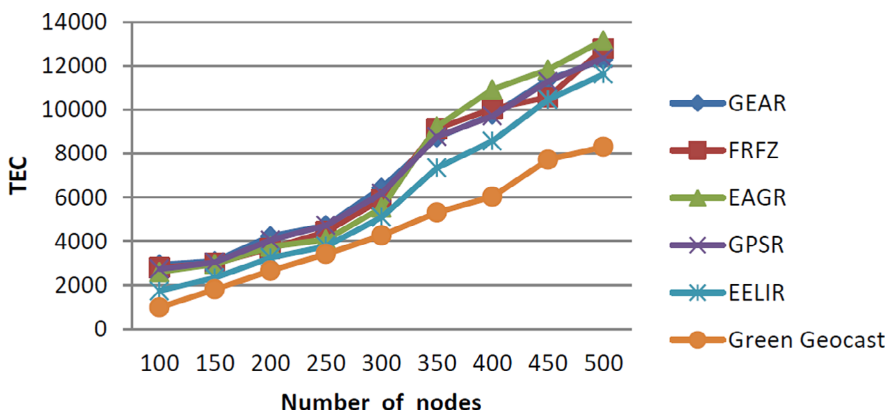

FIGURE 17 Total energy consumption vs number of nodes in the network

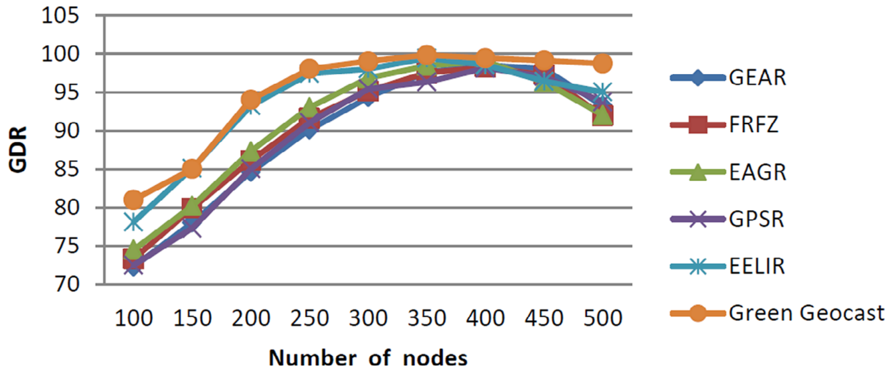

FIGURE 18 Geocast delivery ratio vs number of nodes in the network

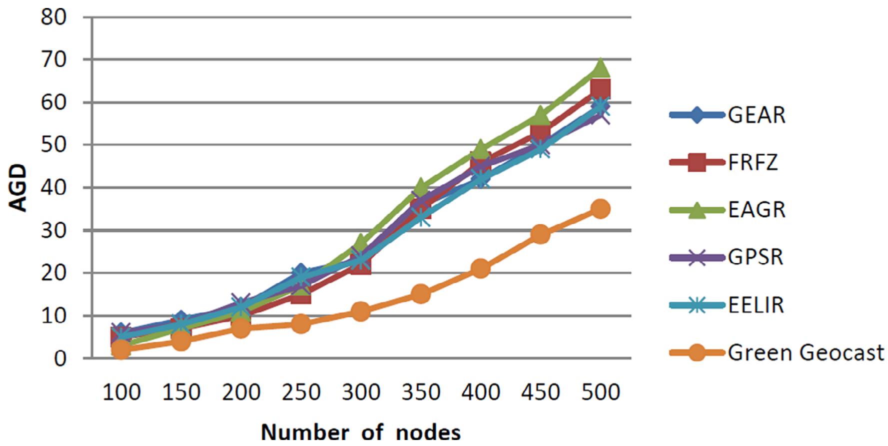

FIGURE 19 Average Geocast Delay vs number of nodes in the network

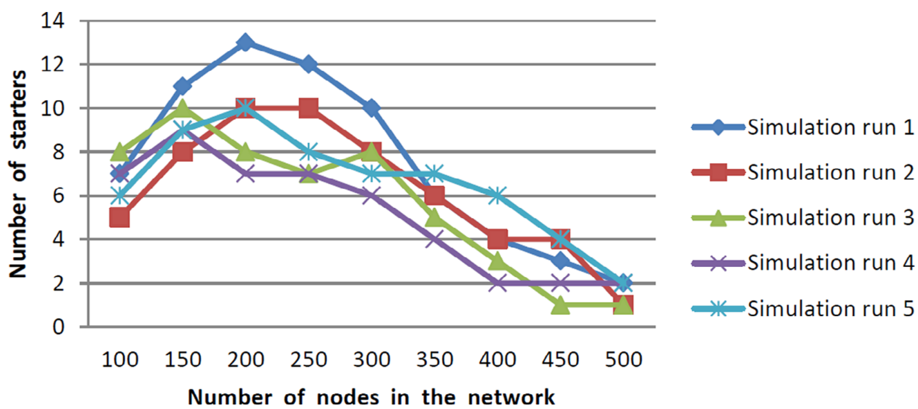

FIGURE 20 Number of starters vs number of nodes in the network

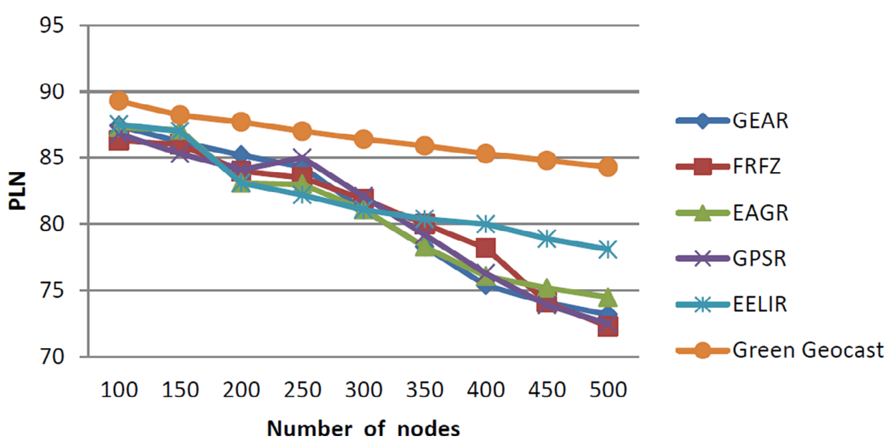

FIGURE 21 Percentage of live nodes vs number of nodes

in geocast region which are reachable from the current router without crossing any edge of the geocast region. Therefore redundancies arise both before entering the geocast region and after entering it. This redundancy is eliminated in our proposed approach because each geocast member receives the broadcast message from exactly one uplink neighbor and before entering the geocast region, geocast message travels from geocast source (if geocast source is inside the current zone) or appropriate incoming peripheral (if geocast source is not inside the current zone) to selected starters, through the most energy efficient path, that is, through only one path. So, there is no redundancy inside portion of geocast region in current zone as well as outside geocast region in the zone. All these save a huge number of messages and therefore a great amount of energy, as revealed in Figures 16 and 17. Lesser number of messages mean lesser contention and collision in the network. So, geocast delivery ratio produced by our proposed approach is much higher than others.

5.5 | Green geocast vs GPSR

GPSR particularly applies perimeter flooding for unicasting. Hence a periphery is created around the geocast region and all nodes in this periphery flood geocast message inside the region. Flooding is always associated to redundancy and redundant transmissions mean huge message cost and wastage of energy. This is evident from Figures 16 and 17. Also, the problem of holes can wreak havoc because nodes in the perimeter deplete quickly due to load of flooding. This increases the risk of breaking perimeter wall producing more holes. As the number of holes increase geocast delivery ratio decreases even more which can be seen in Figure 18.

Green geocast does not require flooding. Messages are transmitted to specific destinations and that also, through minimum energy consuming paths, thereby reducing message cost and energy consumption as shown in Figures 16 and 17. Relocating lonely leaves are performed only when gaining in terms of energy and delay are sure. Therefore our proposed approach gains in terms of delay and successful delivery too.

5.6 | Green geocast vs EELIR

EELIR converts each broadcast operation to multiple unicasts. Although that solves redundancy in broadcast operation but still message cost greatly increases if the network is densely populated and/ or geocast source is significantly far from the geocast region. Green geocast solves this problem by choosing a minimum energy-consuming path to a starter and minimum energy-consuming paths from starter to all nonstarter geocast members. Thus, message transmission/forwarding cost in green geocast is much smaller than EELIR.

EELIR does not rely on building a periphery that holes can wreak havoc. So in low-density networks EELIR produces a better packet delivery ratio than others which is close to green geocast. But as the number of nodes start to increase, message cost and energy consumption in nodes also increase. Along with its message contention and collision in the network increase, reducing geocast delivery ratio.

5.7 | Discussion on number of starters

As seen in Figure 21, the number of starters in initially high and increases a bit with an increase in the number of nodes. This happens when all the newcomer nodes are not reachable from existing starters. But when the number of nodes increases even more then the number of starters begins to decrease; new links are formed joining some existing starters. Therefore the number of starters reduces and also this factor improves the energy efficiency of the network because whenever a starter changes its status to nonstarter geocast member, then the route connecting geocast source or incoming peripheral node to that starter becomes unnecessary and energy consumption through this path is completely saved now.

6 | CONCLUSION

The proposed green geocast is the first one for geocast protocol in a SD-WSN framework. This technique eliminates redundancy in geocast operation and significantly decreases message cost in the network and energy consumption. Hence, message contention and collision have been decreased significantly, leading to substantial enhancement in geocast delivery ratio. Efficient selection of starter along with elevation of lonely leaves significantly improves energy efficiency and also reduces delay in completing the geocast. These are all accomplished by fuzzy controllers that inculcate intelligence into the system.

Geographic addressing and routing have many potential applications in geographic messaging, geographic advertising, delivery of geographically restricted services, and presence discovery of a service or mobile network participant in a limited geographic area. The green geocast method proposed here can be applied in all these cases; the only requirement is that the network is static. As part of future work, the implemention of anycast operation using the starter concept could be invistigated further.

CONFLICTS OF INTEREST

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome. We confirm that the manuscript has been read and approved by all named authors and that there are no other persons who satisfied the criteria for authorship but are not listed. We further confirm that the order of authors listed in the manuscript has been approved by all of us. We confirm that we have given due consideration to the protection of intellectual property associated with this work and that there are no impediments to publication, including the timing of publication, with respect to intellectual property. In so doing we confirm that we have followed the regulations of our institutions concerning intellectual property. We understand that the Corresponding Author is the sole contact for the Editorial process (including Editorial Manager and direct communications with the office). He/she is responsible for communicating with the other authors about progress, submissions of revisions and final approval of proofs.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

ORCID

Ali Safaa Sadiq https://orcid.org/0000-0002-5746-0257

REFERENCES

1. Akyildiz I, Su W, Sankarasubramaniam Y, Cayirci E. Wireless sensornetworks: a survey. *Comput Netw*. 2002;38(4):393-422. https://www.sciencedirect.com/science/article/abs/pii/S1389128601003024.
2. Jin X, Kong FX, Kong LH, et al. A hierarchical data transmission framework for industrial wireless sensor and actuator networks. *IEEE Trans Ind Inf*. 2017;13(4):2019-2029. https://ieeexplore.ieee.org/abstract/document/7885069/.
3. Jan N, Javaid N, Javaid Q, et al. A balanced energy-consuming and hole-alleviating algorithm for wireless sensor networks. *IEEE Access*. 2017;5:6134-6150.
4. Thangaramya K, Kulothungan K, Logambigai R, Selvi M, Ganapathy S, Kannan A. Energy aware cluster and neuro-fuzzy based routing algorithm for wireless sensor networks in IoT. *Comput Netw*. 2019;151:211-223.
5. Ash JN. SeNeLEx: the sensor network localization explorer. http://www.ece.osu.edu/~ashj/localization/
6. Ash JN, Moses RL. Outlier compensation in sensor network self-localization via the EM algorithm. *IEEE ICASSP*. 2005;19-23:278-283. https://ieeexplore.ieee.org/abstract/document/936317/.
7. Ash JN, Potter LC. Sensor network localization via received signal strengthmeasurements with directional antennas. Paper presented at: Proceedings of the 2004 AllertonConference on Communication, Control, and Computing, Illinois, IL; 2004.
8. Sabri Y, Kamoun NL. Geographic routing in wireless sensor networks based on a partitioned architecture. *Int J Comput Appl*. 2016;153(5):975-987.
9. Bulusu N, Heidemann J, Estrin D. GPS-less low cost outdoor localization forvery small devices. *IEEE Personal Comm*. 2000;5(5):28-34.
10. Banerjee A, Sufian A. Smart-Green-Mult (SGM): overhear from topological kingpins in software defined wireless sensor networks. *J Ambient Intell Humaniz Comput*. 2020;12:387-404. https://doi.org/10.1007/s12652-020-01984-2.
11. Cevher V, McClellan JH. Sensor array calibration via tracking with the extended Kalman filter. Paper presented at: Proceedings of the Federal Laboratories Symposium on Advanced Sensors, Salt Lake City, UT; March 2001:51-56.
12. Coates M. Distributed particle filters for sensor networks. *IEEE InformationProcessing in Sensor Networks (IPSN)*; New York, NY: Association for Computing Machinery; 2004:99-107.
13. Preeth SKSL, Dhanalakshmi R, Kumar R, Shakeel PM. An adaptive fuzzy rule based energy efficient clustering and immune-inspired routing protocol for WSN-assisted IoT system. *J Ambient Intell Humaniz Comput*. 2018;1-13. https://link.springer.com/article/10.1007/s12652-018-1154-z.

14. Bhushan B, Sahoo G. Routing protocols in wireless sensor networks. *Computational Intelligence in Sensor Networks*. Berlin/Heidelberg, Germany: Springer; 2019:215-248.
15. Costa JA, Patwari N, Hero AO III. Distributed multidimensional scaling withadaptive weighting for node localization in sensor networks. *IEEE/ACM Trans Sens Netw*. 2006;2(1):39-64. http://www.eecs.umich.edu/~hero/comm.html.
16. Karp B, Kung HT. GPSR: greedy perimeter stateless routing for wireless networks. Paper presented at: Proceedings of the 6th Annual International Conference on Mobile Computing and Networking, Boston, MA; 2000:243-254.
17. Doherty L, Pister KSJ, Ghaoui LE. Convex position estimation in wirelesssensor networks. *IEEE INFOCOM*. Anchorage, AK: IEEE; 2001;3:1655-1663.
18. Kim K, Lee H, Lee B, Baek Y, Han K. A location based energy efficient intersection routing protocol in mobile sensor networks. Paper presented at: Proceedings of the 2008 International Conference on MultiMedia and Information Technology, Three Gorges, China; 2008.
19. Muthumanikandan V, Valliyammai C. Link failure recovery using shortest path fast rerouting technique in SDN. *Wirel Pers Commun*. 2017;97(2):2475-2495.
20. Akila UM. A survey on recent techniques for energy efficient routing in WSN. *Int J Sens Sens Netw*. 2018;6(1):8-15. https://doi.org/10.11648/j.ijssn.20180601.12.
21. Jayarajan P, Kanagachidambaresan GR, Sundararajan TVP, Sakthipandi K, Maheswar R, Karthikeyan A. An energy-aware buffer management (EABM) routing protocol for WSN. *J Supercomput*. 2020;76(6):4543-4555.
22. Leccese, F. “Remote-control system of high efficiency and intelligent street lighting using a zig bee network of devices and sensors”, *IEEE Trans Power Deliv*, 28 (1), 6389795, pp. 21-28, 2013 https://doi.org/10.1109/TPWRD.2012.2212215
23. “Mininet”, [Online]. Available: http://mininet.org/. Accessed November 16, 2017.
24. Guerrero JA, Cosío M, Espinoza A, et al. GeoSoc: a geocast-based communication protocol for monitoring of marine environments. *IEEE Lat Am Trans*. 2017;15(2):324-332.
25. Seada K, Helmy A. Efficient and robust geocasting protocols for sensor networks. *Comput Commun*. 2006;29(2):151-161.
26. Seada K, Helmy A. Geographic protocols in sensor networks. *Encyclopedia of Sensors*. California, CA: American Scientific Publishers (ASP); 2004.
27. Bhushan B, Sahoo G. ISFC-BLS (intelligent and secured fuzzy clustering algorithm using balanced load sub-cluster formation) In WSN environment. *Wirel Pers Commun*. 2020;111:1667-1694. https://doi.org/10.1007/s11277-019-06948-0.
28. Bhushan B, Sahoo G. E2SR2: An acknowledgement-based mobile sink routing protocol with rechargeable sensors for wireless sensor networks. *Wirel Netw*. 2019;25(1):2697-2721.
29. Rajesh G, Raajini XM, Sagayam KM, Bhushan B, Kose U. Fuzzy genetic based dynamic spectrum allocation approach for cognitive radio sensor networks. *Turk J Electr Eng Comput Sci*. 2020;28(5):2416-2432.
30. Selvi M, Kumar SS, Ganapathy S, Ayyanar A, Nehemiah HK, Kannan A. An energy efficient clustered gravitational and fuzzy based routing algorithm in WSNs. *Wirel Pers Commun*. 2021;116:61-90.
31. Younis O, Fahmy S. HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. *IEEE Trans Mob Comput*. 2004;3(4):366-379.
32. Bhushan B, Sahoo G. Recent Advances in Attacks, Technical Challenges, Vulnerabilities and Their Countermeasures in Wireless Sensor Networks. *Wirel Pers Commun*. 2018;98(3):2037-2077.
33. Selvi M, Thangaramya K, Ganapathy S, Kulothungan K, Nehemiah HK, Kannan A. An energy aware trust based secure routing algorithm for effective communication in wireless sensor networks. *Wirel Pers Commun*. 2019;105(4):1475-1490.
34. Satija S, Sharma T, Bhushan B. Innovative approach to wireless sensor networks: sd-wsn. Paper presented at: Proceedings of the International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), Greater Noida, India; 2019.
35. Kadhim AJ, Seno SAH. Energy-efficient multicast routing protocol based on SDN and fog computing for vehicular networks. *Ad Hoc Netw*. 2019;84:68-81.
36. Munuswamy S, Velvizhy P, Ganapathy S, Nehemiah K, Arputharaj K. A rule based delay constrained energy efficient routing technique for wireless sensor networks. *Clust Comput*. 2019;22(5):10839-10848.
37. Sangeetha G, Vijayalakshmi M, Ganapathy S, Kannan A. An improved congestion-aware routing mechanism in sensor networks using fuzzy rule sets. *Peer-to-Peer Netw Appl*. 2020;13(3):890-904.
38. Leccese F, Cagnetti M, Calogero A, et al. A new acquisition and imaging system for environmental measurements: an experience on the Italian cultural heritage. *Sensors (Switzerland)*. 2014;14(5):9290-9312. https://doi.org/10.3390/s140509290.
39. Grover J, Sharma S, Sharma M. Location based protocols in wireless sensor networks - a review. *IEEE ICCCNT*; China: IEEE; 2014.
40. Masoudi R, Ghaffari A. Software defined networks: a survey. *J Netw Comput Appl*. 2016;67:1-25.
41. Moazzeni S, Khayyambashi MR, Movahhedinia N, Callegati F. On reliability improvement of software-defined networks. *Comput Netw*. 2018;133:195-211.
42. “Floodlight”,[Online]. Available: http://www.projectfloodlight.org/floodlight/. Accessed November 16, 2017.
43. Faheem M, Butt RA, Raza B, Ashraf MW, Ngadi MA, Gungor VC. Energy efficient and reliable data gathering using internet of software-defined mobile sinks for WSNs-based smart grid applications. *Comput Stand Interf*. 2019;66:103341.
44. Duan Y, Li W, Fu X, Luo Y, Yang L. A methodology for reliability of WSN based on SDN in adaptive industrial environment. *IEEE/CAA J Automat Sin*. 2018;5(1):74-82.
45. Letswamotse BB, Malekian R, Modieginyane KM. Adaptable QoS provisioning for efficient traffic-to-resource control in software defined wireless sensor networks. *J Ambient Intell Humaniz Comput*. 2020;11(6):2397-2405.

46. Li L, Halpern JY. Minimum-energy mobile wireless networks revisited. Paper presented at: Proceedings of the ICC 2001 IEEE International Conference on Communications Conference Record (Cat. No. 01CH37240), IEEE, Helsinki, Finland; 2001;1:278-283.
47. Yu Y, Govindan R, Estrin D. Geographical and energy aware routing: a recursive data dissemination protocol for wireless sensor networks; 2001.
48. Haider R, Javed MY, Khattak N. EAGR: energy aware greedy routing in sensor networks; January 2008.
49. Kobo HI, Abu-Mahfouz AM, Hancke GP. Efficient controller placement and reflection mechanism in distributed control system for software defined wireless sensor networks. *Trans Emerg Telecommun Technol*. 2019;30(6):1-19.
50. Ghiasian A, Liu P, Wang X, Collier M. Energy efficient operation of a network of OpenFlow switches featuring hardware acceleration and frequency scaling. *Trans Emerg Telecommun Technol*. 2019;30(6):1-15.
51. Huang M, Yu B. LCD: light weight control model for data plane in software defined wireless sensor networks. *Trans Emerg Telecommun Technol*. 2019;30(6):1-17.

How to cite this article: Banerjee A, Sufian A, Sadiq AS, Mirjalili S. Minimum energy transmission forest-based geocast in software-defined wireless sensor networks. *Trans Emerging Tel Tech*. 2021;e4253. https://doi.org/10.1002/ett.4253